GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 8, 2002, 11:43:04 ; Search time 39.63 Seconds
(without alignments)
52.383 Million cell updates/sec

Title: US-09-461-061A-2
Perfect score: 60
Sequence: 1 TLTHTITKLNAE 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 1842

Minimum DB seq length: 0
Maximum DB seq length: 12

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	23	38.3	8	2	Q45615	Q45615 bacillus su
2	22	36.7	11	8	O79921	O79921 phrynocepha
3	20.5	34.2	10	6	Q9TRU6	Q9tru6 bos taurus
4	18	30.0	7	2	Q54248	Q54248 streptomyce
5	18	30.0	8	3	Q05403	Q05403 saccharomyc
6	18	30.0	9	4	Q14277	Q14277 homo sapien
7	18	30.0	10	6	Q9XS84	Q9xs84 equus cabal
8	18	30.0	11	5	Q9VN99	Q9vn99 drosophila
9	17	28.3	8	2	P83158	P83158 anabaena sp
10	17	28.3	10	11	Q91WZ3	Q91wz3 rattus sp.
11	17	28.3	11	8	Q94VK1	Q94vk1 varanus aca
12	17	28.3	11	8	Q94VG8	Q94vg8 varanus gou
13	17	28.3	11	8	Q94VE7	Q94ve7 varanus kom
14	17	28.3	11	8	Q94VB8	Q94vb8 varanus sal
15	17	28.3	11	9	Q37925	Q37925 bacteriopha
16	17	28.3	12	2	Q50342	Q50342 mycoplasma
17	17	28.3	12	11	Q9WUX1	Q9wux1 rattus norv
18	17	28.3	12	16	Q9K670	Q9k670 bacillus ha
19	16	26.7	7	2	P72081	P72081 nocardia la
20	16	26.7	8	4	Q15889	Q15889 homo sapien
21	16	26.7	9	4	Q9H4M8	Q9h4m8 homo sapien
22	16	26.7	9	12	P90359	P90359 barley mild
23	16	26.7	10	8	Q9TGA1	Q9tga1 heloderma s
24	16	26.7	10	8	Q9TG98	Q9tg98 shinisaurus
25	16	26.7	10	8	Q9TG83	Q9tg83 diploglossu
26	16	26.7	10	8	Q9TG77	Q9tg77 sauresia ag
27	16	26.7	10	8	Q9TG74	Q9tg74 wetmorena h
28	16	26.7	10	8	Q9TG44	Q9tg44 anguis frag
29	16	26.7	10	8	P92774	P92774 xantusia vi
30	16	26.7	10	8	P92616	P92616 cnemidophor
31	16	26.7	10	8	P92766	P92766 varanus gri
32	16	26.7	10	8	Q9TG92	Q9tg92 anniella pu
33	16	26.7	11	7	O77884	O77884 oreochromis
34	16	26.7	11	7	O77885	O77885 oreochromis
35	16	26.7	11	8	Q94V77	Q94v77 heloderma s
36	16	26.7	12	2	P83054	P83054 bacteroides
37	16	26.7	12	3	Q9UR22	Q9ur22 cryptococcu
38	16	26.7	12	15	Q85498	Q85498 equine infe
39	15	25.0	7	2	Q47029	Q47029 enterobacte
40	15	25.0	8	2	P77556	P77556 escherichia
41	15	25.0	8	6	Q9TT78	Q9tt78 canis famil
42	15	25.0	9	2	Q47556	Q47556 escherichia
43	15	25.0	9	2	Q51349	Q51349 pseudomonas
44	15	25.0	10	2	Q9F5W1	Q9f5w1 vibrio chol
45	15	25.0	10	5	P82223	P82223 bombyx mori

ALIGNMENTS

RESULT 1
Q45615
ID Q45615 PRELIMINARY; PRT; 8 AA.
AC Q45615;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GUTB PROTEIN (FRAGMENT).
GN GUTB.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=94253000; PubMed=8195086;
RA Ye R., Wong S.L.;
RT "Transcriptional regulation of the Bacillus subtilis glucitol
RT dehydrogenase gene.";
RL J. Bacteriol. 176:3314-3320(1994).
DR EMBL; L16626; AAA20875.1; -.
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 927 MW; FD56C772D1A1F1A6 CRC64;

Query Match 38.3%; Score 23; DB 2; Length 8;
Best Local Similarity 60.0%; Pred. No. 5.6e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy      2 LTHTI 6
          :|||:
Db      1 MTHTV 5
```

RESULT 2
O79921
ID O79921 PRELIMINARY; PRT; 11 AA.
AC O79921;

```
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   CYTOCHROME C OXIDASE SUBUNIT I (FRAGMENT).
GN   COI.
OS   Phrynocephalus raddei.
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Iguania; Acrodonta; Agamidae; Agaminae;
OC   Phrynocephalus.
OX   NCBI_TaxID=52206;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97315309; PubMed=9169559;
RA   Macey J.R., Larson A., Ananjeva N.B., Papenfuss T.J.;
RT   "Evolutionary shifts in three major structural features of the
RT   mitochondrial genome among iguanian lizards.";
RL   J. Mol. Evol. 44:660-674(1997).
DR   EMBL; U82691; AAC62302.1; -.
KW   Mitochondrion.
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  1324 MW;  5DF73710D371F1A7 CRC64;


  Query Match             36.7%;  Score 22;  DB 8;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 TLTH 4
          ||||
Db      3 TLTH 6


RESULT   3
Q9TRU6
ID   Q9TRU6      PRELIMINARY;      PRT;    10 AA.
AC   Q9TRU6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   GAP-3, GTPASE-ACTIVATING PROTEIN.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92112868; PubMed=1309786;
RA   Nice E.C., Fabri L., Hammacher A., Holden J., Simpson R.J.,
RA   Burgess A.W.;
RT   "The purification of a Rap1 GTPase-activating protein from bovine
RT   brain cytosol.";
RL   J. Biol. Chem. 267:1546-1553(1992).
SQ   SEQUENCE   10 AA;  1149 MW;  59370A51A72321A7 CRC64;


  Query Match             34.2%;  Score 20.5;  DB 6;  Length 10;
  Best Local Similarity   75.0%;  Pred. No. 2e+03;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    1;  Gaps    1;

Qy      6 ITKL-NAE 12
          :||| |||
Db      3 LTKLTNAE 10


RESULT   4
Q54248
ID   Q54248      PRELIMINARY;      PRT;     7 AA.
AC   Q54248;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   RPLO PROTEIN (FRAGMENT).
GN   RPLO.
OS   Streptomyces griseus.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1911;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N2-3-11;
RX   MEDLINE=20011291; PubMed=10542330;
RA   Poehling S., Piepersberg W., Wehmeier U.F.;
RT   "Analysis and regulation of the sec Y gene from Streptomyces griseus
RT   N2-3-11 and interaction of the SecY protein with the SecA protein.";
RL   Biochim. Biophys. Acta 1447:298-302(1999).
DR   EMBL; X95915; CAA65160.1; -.
FT   NON_TER      1      1
SQ   SEQUENCE   7 AA;  760 MW;  72C72B01B2D1B2A0 CRC64;


  Query Match             30.0%;  Score 18;  DB 2;  Length 7;
  Best Local Similarity   60.0%;  Pred. No. 5.6e+05;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 TITKL 9
          |:|:|
Db      2 TVTEL 6


RESULT   5
Q05403
ID   Q05403      PRELIMINARY;      PRT;     8 AA.
AC   Q05403;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   DNA FOR ORF'S FROM CHROMOSOME XV (FRAGMENT).
GN   COQ3 AND YOL096C.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FY1679;
RX   MEDLINE=96021609; PubMed=8533473;
RA   Zumstein E., Pearson B.M., Kalogeropoulos A., Schweizer M.;
RT   "A 29.425 kb segment on the left arm of yeast chromosome XV contains
RT   more than twice as many unknown as known open reading frames.";
RL   Yeast 11:975-986(1995).
DR   EMBL; X83121; CAA58183.1; -.
DR   SGD; S0005456; COQ3.
FT   NON_TER      8      8
SQ   SEQUENCE   8 AA;  879 MW;  7B5322D2C441E058 CRC64;


  Query Match             30.0%;  Score 18;  DB 3;  Length 8;
  Best Local Similarity   33.3%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 HTITKL 9
          | : |:
Db      3 HNVVKI 8


RESULT   6
Q14277
ID   Q14277      PRELIMINARY;      PRT;     9 AA.
AC   Q14277;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   RET PROTEIN SHORT FORM (FRAGMENT).
OS   Homo sapiens (Human).
```

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94071887; PubMed=7902707;
RA   Ceccherini I., Bocciardi R., Luo Y., Pasini B., Hofstra R.,
RA   Takahashi M., Romeo G.;
RT   "Exon structure and flanking intronic sequences of the human RET
RT   proto-oncogene.";
RL   Biochem. Biophys. Res. Commun. 196:1288-1295(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94366753; PubMed=8084609;
RA   Ceccherini I., Hofstra R., Yin L., Stulp R., Barone V., Stelwagen T.,
RA   Bocciardi R., Nijveen H., Bolino A., Seri M., Ronchetto P., Pasini B.,
RA   Bozzano M., Buys C., Romeo G.;
RT   "DNA polymorphisms and conditions for SSCP analysis of the 20 exons of
RT   the ret proto-oncogene.";
RL   Oncogene 9:3025-3029(1994).
DR   EMBL; U11532; AAC50102.1; -.
FT   NON_TER      1      1
SQ   SEQUENCE   9 AA;  1134 MW;  034E11A9CDD1F5B0 CRC64;


  Query Match            30.0%;  Score 18;  DB 4;  Length 9;
  Best Local Similarity  28.6%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 LTHTITK 8
           ::|  |:
Db       2 ISHAFTR 8


RESULT   7
Q9XS84
ID   Q9XS84      PRELIMINARY;      PRT;    10 AA.
AC   Q9XS84;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   LEPTIN (FRAGMENT).
GN   LEP.
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID=9796;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99160468; PubMed=10051323;
RA   Caetano A.R., Pomp D., Murray J.D., Bowling A.T.;
RT   "Comparative mapping of 18 equine type I genes assigned by somatic
RT   cell hybrid analysis.";
RL   Mamm. Genome 10:271-276(1999).
DR   EMBL; AF097582; AAD25985.1; -.
FT   NON_TER      1      1
FT   NON_TER     10     10
SQ   SEQUENCE   10 AA;  1155 MW;  6BE77BB05AA44044 CRC64;


  Query Match            30.0%;  Score 18;  DB 6;  Length 10;
  Best Local Similarity  40.0%;  Pred. No. 6.2e+03;
  Matches    2;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 ITKLN 10
           :|::|
Db       1 VTRIN 5


RESULT   8
Q9VN99
ID   Q9VN99      PRELIMINARY;      PRT;    11 AA.
AC   Q9VN99;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   CG2676 PROTEIN.
GN   CG2676.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003603; AAF52049.1; -.
DR   FlyBase; FBgn0037309; CG2676.
SQ   SEQUENCE   11 AA;  1113 MW;  8760FE9BD5A1B1B1 CRC64;


  Query Match            30.0%;  Score 18;  DB 5;  Length 11;
  Best Local Similarity  27.3%;  Pred. No. 6.8e+03;
  Matches    3;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 LTHTITKLNAE 12
           :| | | ::::
Db       1 MTATTTSVSSK 11


RESULT   9
P83158
ID   P83158      PRELIMINARY;      PRT;     8 AA.
AC   P83158;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
```

```
DE   PHOTOSYSTEM I IRON-SULFUR CENTER (PHOTOSYSTEM I SUBUNIT VII) (9 KDA
DE   POLYPEPTIDE) (PSI-C) (FRAGMENT).
OS   Anabaena sp. (strain L31).
OC   Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.
OX   NCBI_TaxID=29412;
RN   [1]
RP   SEQUENCE.
RA   Apte S.K., Uhlemann E., Schmid R., Altendorf K.;
RL   Submitted (OCT-2001) to the SWISS-PROT data bank.
CC   -!- FUNCTION: APOPROTEIN FOR THE IRON-SULFUR CENTERS FA AND FB OF THE
CC       PHOTOSYSTEM I COMPLEX.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF
CC       'BACTERIAL-TYPE' 4FE-4S FERREDOXINS.
DR   InterPro; IPR001450; 4Fe4S_ferredoxin.
DR   PROSITE; PS00198; 4FE4S_FERREDOXIN; PARTIAL.
KW   Photosynthesis; Photosystem I; Iron-sulfur; 4Fe-4S; Metal-binding.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  962 MW;  C5BB505322D1A1F5 CRC64;


  Query Match             28.3%;  Score 17;  DB 2;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 THTI 6
          :||:
Db      1 SHTV 4


RESULT  10
Q91WZ3
ID   Q91WZ3      PRELIMINARY;      PRT;    10 AA.
AC   Q91WZ3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   LUTEINIZING HORMONE/CHORIONIC GONADOTROPIN RECEPTOR HOMOLOG
DE   (FRAGMENT).
OS   Rattus sp.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10118;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=OVARY;
RX   MEDLINE=96147985; PubMed=8571710;
RA   Shen Q.X., Liu H.H., Chen W.Y., Bahl O.P.;
RT   "[Cloning and overexpression of rat ovary LH/hCG receptor cDNA in
RT   insect cells].";
RL   Shih Yen Sheng Wu Hsueh Pao 28:283-290(1995).
DR   EMBL; S80660; AAB50710.1; -.
KW   Receptor; Chorion.
FT   NON_TER       1      1
SQ   SEQUENCE   10 AA;  1129 MW;  09A5F22DC4177760 CRC64;


  Query Match             28.3%;  Score 17;  DB 11;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 9.6e+03;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 LTH 4
          |||
Db      8 LTH 10


RESULT  11
Q94VK1
ID   Q94VK1      PRELIMINARY;      PRT;    11 AA.
AC   Q94VK1;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CYTOCHROME C OXIDASE SUBUNIT I (FRAGMENT).
GN   COI.
OS   Varanus acanthurus.
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
OX   NCBI_TaxID=62035;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ast J.C.;
RT   "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
RL   Cladistics 17:0-0(2001).
DR   EMBL; AF407488; AAL10021.1; -.
KW   Mitochondrion.
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  1370 MW;  8E6DEE80C7336411 CRC64;


  Query Match             28.3%;  Score 17;  DB 8;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 1.1e+04;
  Matches    4;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 TLTHTITKLN 10
          |||  :   |
Db      2 TLTRWLFSTN 11


RESULT  12
Q94VG8
ID   Q94VG8      PRELIMINARY;      PRT;    11 AA.
AC   Q94VG8;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CYTOCHROME C OXIDASE SUBUNIT I (FRAGMENT).
GN   COI.
OS   Varanus gouldii.
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
OX   NCBI_TaxID=62042;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ast J.C.;
RT   "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
RL   Cladistics 17:0-0(2001).
DR   EMBL; AF407502; AAL10060.1; -.
KW   Mitochondrion.
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  1370 MW;  8E6DEE80C7336411 CRC64;


  Query Match             28.3%;  Score 17;  DB 8;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 1.1e+04;
  Matches    4;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 TLTHTITKLN 10
          |||  :   |
Db      2 TLTRWLFSTN 11


RESULT  13
Q94VE7
ID   Q94VE7      PRELIMINARY;      PRT;    11 AA.
AC   Q94VE7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CYTOCHROME C OXIDASE SUBUNIT I (FRAGMENT).
GN   COI.
OS   Varanus komodoensis.
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC   Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
OX   NCBI_TaxID=61221;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ast J.C.;
RT   "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
RL   Cladistics 17:0-0(2001).
DR   EMBL; AF407510; AAL10084.1; -.
KW   Mitochondrion.
FT   NON_TER      11      11
SQ   SEQUENCE   11 AA;  1370 MW;  8E6DEE80C7336411 CRC64;

  Query Match             28.3%;  Score 17;  DB 8;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 1.1e+04;
  Matches    4;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKLN 10
           |||  :   |
Db       2 TLTRWLFSTN 11


RESULT  14
Q94VB8
ID   Q94VB8       PRELIMINARY;       PRT;    11 AA.
AC   Q94VB8;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CYTOCHROME C OXIDASE SUBUNIT I (FRAGMENT).
GN   COI.
OS   Varanus salvadorii.
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
OX   NCBI_TaxID=62049;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ast J.C.;
RT   "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
RL   Cladistics 17:0-0(2001).
DR   EMBL; AF407522; AAL10119.1; -.
KW   Mitochondrion.
FT   NON_TER      11      11
SQ   SEQUENCE   11 AA;  1370 MW;  8E6DEE80C7336411 CRC64;

  Query Match             28.3%;  Score 17;  DB 8;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 1.1e+04;
  Matches    4;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKLN 10
           |||  :   |
Db       2 TLTRWLFSTN 11


RESULT  15
Q37925
ID   Q37925       PRELIMINARY;       PRT;    11 AA.
AC   Q37925;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   BACTERIOPHAGE FR REPLICASE (FRAGMENT).
OS   Bacteriophage fr.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Leviviridae;
OC   Levivirus.
OX   NCBI_TaxID=12017;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Berzin V.M., Gribanov V.A., Cielens I.E., Jansone I.V., Gren E.J.;
RT   "The nucleotide sequence of the regulatory region of phage fr
RT   replicase cistron.";
RL   Bioorg. Khim. 7:306-308(1981).
DR   EMBL; M34834; AAA32193.1; -.
FT   NON_TER      11      11
SQ   SEQUENCE   11 AA;  1285 MW;  8BD43470C33321B1 CRC64;

  Query Match             28.3%;  Score 17;  DB 9;  Length 11;
  Best Local Similarity   30.0%;  Pred. No. 1.1e+04;
  Matches    3;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 LTHTITKLNA 11
           :: |   | |:
Db       1 MSKTTKKFNS 10


RESULT  16
Q50342
ID   Q50342       PRELIMINARY;       PRT;    12 AA.
AC   Q50342;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ATP-SYNTHASE ALPHA (FRAGMENT).
OS   Mycoplasma pneumoniae.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC   Mycoplasmataceae; Mycoplasma.
OX   NCBI_TaxID=2104;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=M129;
RX   MEDLINE=95075318; PubMed=7984111;
RA   Proft T., Herrmann R.;
RT   "Identification and characterization of hitherto unknown Mycoplasma
RT   pneumoniae proteins.";
RL   Mol. Microbiol. 13:337-348(1994).
DR   EMBL; Z32665; CAA83583.1; -.
FT   NON_TER      12      12
SQ   SEQUENCE   12 AA;  1396 MW;  8A7DE4E61CA1A681 CRC64;

  Query Match             28.3%;  Score 17;  DB 2;  Length 12;
  Best Local Similarity   80.0%;  Pred. No. 1.2e+04;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TLTHT 5
           ||| |
Db       3 TLTMT 7


RESULT  17
Q9WUX1
ID   Q9WUX1       PRELIMINARY;       PRT;    12 AA.
AC   Q9WUX1;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   APOBEC-1 PROTEIN (FRAGMENT).
GN   APOBEC-1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98335789; PubMed=9672068;
RA   Greeve J., Axelos D., Welker S., Schipper M., Greten H.;
RT   "Distinct promoters induce APOBEC-1 expression in rat liver and
RT   intestine.";
RL   Arterioscler. Thromb. Vasc. Biol. 18:1079-1092(1998).
DR   EMBL; AJ006695; CAB44439.1; -.
FT   NON_TER      1       1
FT   NON_TER      12      12
SQ   SEQUENCE   12 AA;  1357 MW;  70FB1679699325BB CRC64;
```

```
Query Match             28.3%;  Score 17;  DB 11;  Length 12;
Best Local Similarity   37.5%;  Pred. No. 1.2e+04;
Matches    3;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 TITKLNAE 12
           | :|:::|
Db       4 TESKMSSE 11


RESULT  18
Q9K670
ID   Q9K670      PRELIMINARY;      PRT;    12 AA.
AC   Q9K670;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   BH3862 PROTEIN.
GN   BH3862.
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C-125 / JCM 9153;
RX   MEDLINE=20512582; PubMed=11058132;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
DR   EMBL; AP001520; BAB07581.1; -.
KW   Complete proteome.
SQ   SEQUENCE   12 AA;  1463 MW;  44E501F9695321A7 CRC64;


  Query Match             28.3%;  Score 17;  DB 16;  Length 12;
  Best Local Similarity   25.0%;  Pred. No. 1.2e+04;
  Matches    2;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 LTHTITKL 9
           : : :|||:
Db       1 MNYLLTKI 8


RESULT  19
P72081
ID   P72081      PRELIMINARY;      PRT;     7 AA.
AC   P72081;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   3'-METHYLCEPHEM HYDROXYLASE (FRAGMENT).
GN   CEFF.
OS   Nocardia lactamdurans.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Pseudonocardineae; Pseudonocardiaceae; Amycolatopsis.
OX   NCBI_TaxID=1913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96009872; PubMed=7557411;
RA   Coque J., Perez-Llarena F.J., Enguita F.J., Fuente J.L., Martin J.F.,
RA   Liras P.;
RT   "Characterization of the cmcH genes of Nocardia lactamdurans and
RT   Streptomyces clavuligerus encoding a functional 3'-hydroxymethylcephem
RT   O-carbamoyltransferase for cephamycin biosynthesis.";
RL   Gene 162:21-27(1995).
DR   EMBL; Z21682; CAA79797.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   7 AA;  746 MW;  75A1B2CDD1E681C0 CRC64;


  Query Match             26.7%;  Score 16;  DB 2;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 HTIT 7
           | :|
Db       3 HAVT 6


RESULT  20
Q15889
ID   Q15889      PRELIMINARY;      PRT;     8 AA.
AC   Q15889;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   (CLONE XP15H8B) (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA;
RA   Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA   Coolbaugh M.I., Chinault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,
RA   Caskey C.T.H.;
RT   "Isolation of chromosome-specific genes by reciprocal probing of
RT   arrayed cDNAs and cosmid libraries.";
RL   Hum. Mol. Genet. 0:0-0(1995).
DR   EMBL; L32070; AAA73879.1; -.
FT   NON_TER       1       1
FT   NON_TER       8       8
SQ   SEQUENCE   8 AA;  865 MW;  0474472325A761E7 CRC64;


  Query Match             26.7%;  Score 16;  DB 4;  Length 8;
  Best Local Similarity   75.0%;  Pred. No. 5.6e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 TKLN 10
           :|||
Db       4 SKLN 7


Search completed: July  8, 2002, 11:48:39
Job time: 335 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 8, 2002, 11:44:04 ; Search time 12.97 Seconds
(without alignments)
60.264 Million cell updates/sec

Title: US-09-461-061A-4
Perfect score: 161
Sequence: 1 TLTHTITKLNAENNATFYFKIDNVKKARVQVV 32

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 146970

Minimum DB seq length: 0
Maximum DB seq length: 32

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	95	59.0	26	4	US-08-676-242-15	Sequence 15, Appl
2	41	25.5	32	1	US-07-946-421-17	Sequence 17, Appl
3	39	24.2	31	2	US-08-318-157B-28	Sequence 28, Appl
4	39	24.2	32	1	US-07-946-421-9	Sequence 9, Appli
5	39	24.2	32	1	US-08-137-117D-122	Sequence 122, App
6	39	24.2	32	1	US-08-137-117D-124	Sequence 124, App
7	39	24.2	32	1	US-08-137-117D-141	Sequence 141, App
8	39	24.2	32	2	US-08-436-717-122	Sequence 122, App
9	39	24.2	32	2	US-08-436-717-124	Sequence 124, App
10	39	24.2	32	2	US-08-436-717-141	Sequence 141, App
11	39	24.2	32	2	US-08-553-501A-67	Sequence 67, Appl
12	39	24.2	32	2	US-08-553-501A-71	Sequence 71, Appl
13	39	24.2	32	2	US-08-560-558E-3	Sequence 3, Appli
14	39	24.2	32	2	US-08-765-783A-89	Sequence 89, Appl
15	39	24.2	32	2	US-08-765-783A-91	Sequence 91, Appl
16	39	24.2	32	3	US-09-205-231-67	Sequence 67, Appl
17	39	24.2	32	3	US-09-205-231-71	Sequence 71, Appl
18	39	24.2	32	4	US-08-646-265A-123	Sequence 123, App
19	39	24.2	32	4	US-09-416-557-89	Sequence 89, Appl
20	39	24.2	32	4	US-09-416-557-91	Sequence 91, Appl
21	37	23.0	20	2	US-08-053-451B-169	Sequence 169, App
22	37	23.0	32	2	US-08-470-139-13	Sequence 13, Appl
23	37	23.0	32	4	US-08-569-147-40	Sequence 40, Appl
24	37	23.0	32	4	US-09-347-061-13	Sequence 13, Appl
25	37	23.0	32	4	US-09-425-638A-92	Sequence 92, Appl
26	37	23.0	32	4	US-09-425-638A-93	Sequence 93, Appl
27	37	23.0	32	4	US-09-425-638A-94	Sequence 94, Appl
28	37	23.0	32	4	US-09-543-004-92	Sequence 92, Appl
29	37	23.0	32	4	US-09-543-004-93	Sequence 93, Appl
30	37	23.0	32	4	US-09-543-004-94	Sequence 94, Appl
31	36	22.4	18	2	US-08-480-190-239	Sequence 239, App
32	36	22.4	18	2	US-08-488-379-239	Sequence 239, App
33	36	22.4	18	5	PCT-US93-07545-239	Sequence 239, App
34	35	21.7	7	4	US-08-676-242-21	Sequence 21, Appl
35	35	21.7	7	4	US-09-402-732-8	Sequence 8, Appli
36	35	21.7	32	4	US-09-425-638A-96	Sequence 96, Appl
37	35	21.7	32	4	US-09-543-004-96	Sequence 96, Appl
38	34	21.1	25	5	PCT-US91-02942-26	Sequence 26, Appl
39	34	21.1	28	4	US-09-082-279B-873	Sequence 873, App
40	34	21.1	28	4	US-09-315-304B-873	Sequence 873, App
41	34	21.1	32	2	US-08-318-157B-37	Sequence 37, Appl
42	34	21.1	32	4	US-09-425-638A-95	Sequence 95, Appl
43	34	21.1	32	4	US-09-543-004-95	Sequence 95, Appl
44	33	20.5	23	1	US-07-791-930C-10	Sequence 10, Appl
45	33	20.5	23	1	US-08-173-515B-13	Sequence 13, Appl

ALIGNMENTS

```
RESULT   1
US-08-676-242-15
; Sequence 15, Application US/08676242C
; Patent No. 6143719
; GENERAL INFORMATION:
;  APPLICANT: The Regents of the University of Michigan
;  APPLICANT: Schmaier, Alvin H.
;  APPLICANT: Hasan, Ahmed A.K.
;  TITLE OF INVENTION: Bradykinin Analogs As Selective Thrombin Inhibitors
;  FILE REFERENCE: 8820-2 US
;  CURRENT APPLICATION NUMBER: US/08/676,242C
;  CURRENT FILING DATE: 2000-07-16
;  EARLIER APPLICATION NUMBER: 60/000,096
;  EARLIER FILING DATE: 1995-06-09
;  EARLIER APPLICATION NUMBER: PCT/US96/09940
;  EARLIER FILING DATE: 1996-06-07
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
;   LENGTH: 26
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Bradykinin
;   OTHER INFORMATION: analog
US-08-676-242-15


  Query Match             59.0%;  Score 95;  DB 4;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 9.1e-07;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      14 NATFYFKIDNVKKARVQVV 32
           |||||||||||||||||||
Db       1 NATFYFKIDNVKKARVQVV 19


RESULT   2
US-07-946-421-17
; Sequence 17, Application US/07946421
; Patent No. 5558864
;  GENERAL INFORMATION:
;    APPLICANT:  Bendig, Mary M.
;    APPLICANT:  Kettleborough, Catherine A.
;    APPLICANT:  Saldanha, Jose
;    TITLE OF INVENTION:  Humanized and Chimeric Monoclonal
;    TITLE OF INVENTION:  Antibodies
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Millen, White, Zelano & Branigan, P.C.
```

```
;       STREET:  2200 Clarendon Boulevard, Suite 1400
;       CITY:  Arlington
;       STATE:  Virginia
;       COUNTRY:  U.S.A.
;       ZIP:  22201
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/946,421
;       FILING DATE:  06-NOV-1992
;       CLASSIFICATION:  424
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  WO PCT/EP92/00480
;       FILING DATE:  04-MAR-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  EP 911933892
;       FILING DATE:  06-MAR-1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Hamlet-King, Diana
;       REGISTRATION NUMBER:  33,302
;       REFERENCE/DOCKET NUMBER:  Merck 1430
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  703-243-6333
;       TELEFAX:  703-243-6410
;       TELEX:  64191
;   INFORMATION FOR SEQ ID NO:  17:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  32 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
US-07-946-421-17


  Query Match             25.5%;  Score 41;  DB 1;  Length 32;
  Best Local Similarity   50.0%;  Pred. No. 25;
  Matches    7;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 TITKLNAENNATFY 18
           ||::: ||: ||:|
Db      18 TISRMEAEDAATYY 31


RESULT   3
US-08-318-157B-28
; Sequence 28, Application US/08318157B
; Patent No. 5874540
;  GENERAL INFORMATION:
;    APPLICANT:  HANSEN, Hans J.
;    APPLICANT:  ARMOUR, Kathryn L.
;    TITLE OF INVENTION:  CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
;    TITLE OF INVENTION:  MOUSE MONOCLONAL ANTIBODIES
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/318,157B
;      FILING DATE:  05-OCT-1994
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  SAXE, Bernhard D.
;      REGISTRATION NUMBER:  28,665
;      REFERENCE/DOCKET NUMBER:  18733/464
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  4
;      OTHER INFORMATION:  /note= "At site 4, Xaa = Ser or
;      OTHER INFORMATION:  Asp."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  9
;      OTHER INFORMATION:  /note= "At site 9, Xaa = Gly or
;      OTHER INFORMATION:  Val."
US-08-318-157B-28


  Query Match             24.2%;  Score 39;  DB 2;  Length 31;
  Best Local Similarity   50.0%;  Pred. No. 46;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 THTITKLNAENNATFY 18
           | ||: |  |: ||:|
Db      15 TFTISSLQPEDIATYY 30


RESULT   4
US-07-946-421-9
; Sequence 9, Application US/07946421
; Patent No. 5558864
;  GENERAL INFORMATION:
;    APPLICANT:  Bendig, Mary M.
;    APPLICANT:  Kettleborough, Catherine A.
;    APPLICANT:  Saldanha, Jose
;    TITLE OF INVENTION:  Humanized and Chimeric Monoclonal
;    TITLE OF INVENTION:  Antibodies
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Millen, White, Zelano & Branigan, P.C.
;      STREET:  2200 Clarendon Boulevard, Suite 1400
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22201
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/946,421
;      FILING DATE:  06-NOV-1992
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/EP92/00480
;      FILING DATE:  04-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  EP 911933892
;      FILING DATE:  06-MAR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hamlet-King, Diana
;      REGISTRATION NUMBER:  33,302
;      REFERENCE/DOCKET NUMBER:  Merck 1430
;    TELECOMMUNICATION INFORMATION:
```

```
;      TELEPHONE:  703-243-6333
;      TELEFAX:  703-243-6410
;      TELEX:  64191
;   INFORMATION FOR SEQ ID NO:  9:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  32 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     FEATURE:
;       NAME/KEY:  Region
;       LOCATION:  15
;       OTHER INFORMATION:  /note= "Amino acid 15 can be Tyr,
;       OTHER INFORMATION:  Phe, Trp or His."
US-07-946-421-9


  Query Match             24.2%;  Score 39;  DB 1;  Length 32;
  Best Local Similarity   50.0%;  Pred. No. 48;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 THTITKLNAENNATFY 18
           | ||: |  |: ||:|
Db      16 TFTISSLQPEDIATYY 31


RESULT   5
US-08-137-117D-122
; Sequence 122, Application US/08137117D
; Patent No. 5795965
;  GENERAL INFORMATION:
;    APPLICANT:  TSUCHIYA, Masayuki
;    APPLICANT:  SATO, Koh
;    APPLICANT:  BENDIG, Mary
;    APPLICANT:  JONES, Steven
;    APPLICANT:  SALDANHA, Jose
;    TITLE OF INVENTION:  RESHAPED HUMAN ANTIBODY TO HUMAN
;    TITLE OF INVENTION:  INTERLEUKIN-6 RECEPTOR
;    NUMBER OF SEQUENCES:  158
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/137,117D
;      FILING DATE:  20-DEC-1993
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/JP92/00544
;      FILING DATE:  24-APR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 4-32084
;      FILING DATE:  19-FEB-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 3-95476
;      FILING DATE:  25-APR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  WEGNER, Harold C.
;      REGISTRATION NUMBER:  25,258
;      REFERENCE/DOCKET NUMBER:  53466/126/AAOK
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  122:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  32 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-137-117D-122


  Query Match             24.2%;  Score 39;  DB 1;  Length 32;
  Best Local Similarity   50.0%;  Pred. No. 48;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 THTITKLNAENNATFY 18
           | ||: |  |: ||:|
Db      16 TFTISSLQPEDIATYY 31


RESULT   6
US-08-137-117D-124
; Sequence 124, Application US/08137117D
; Patent No. 5795965
;  GENERAL INFORMATION:
;    APPLICANT:  TSUCHIYA, Masayuki
;    APPLICANT:  SATO, Koh
;    APPLICANT:  BENDIG, Mary
;    APPLICANT:  JONES, Steven
;    APPLICANT:  SALDANHA, Jose
;    TITLE OF INVENTION:  RESHAPED HUMAN ANTIBODY TO HUMAN
;    TITLE OF INVENTION:  INTERLEUKIN-6 RECEPTOR
;    NUMBER OF SEQUENCES:  158
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/137,117D
;      FILING DATE:  20-DEC-1993
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/JP92/00544
;      FILING DATE:  24-APR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 4-32084
;      FILING DATE:  19-FEB-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 3-95476
;      FILING DATE:  25-APR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  WEGNER, Harold C.
;      REGISTRATION NUMBER:  25,258
;      REFERENCE/DOCKET NUMBER:  53466/126/AAOK
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  124:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  32 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-137-117D-124


  Query Match             24.2%;  Score 39;  DB 1;  Length 32;
  Best Local Similarity   50.0%;  Pred. No. 48;
```

Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```
Qy      3 THTITKLNAENNATFY 18
          | ||: |  |: ||:|
Db     16 TFTISSLQPEDIATYY 31
```

RESULT 7
US-08-137-117D-141
; Sequence 141, Application US/08137117D
; Patent No. 5795965
; GENERAL INFORMATION:
; APPLICANT: TSUCHIYA, Masayuki
; APPLICANT: SATO, Koh
; APPLICANT: BENDIG, Mary
; APPLICANT: JONES, Steven
; APPLICANT: SALDANHA, Jose
; TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN
; TITLE OF INVENTION: INTERLEUKIN-6 RECEPTOR
; NUMBER OF SEQUENCES: 158
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/137,117D
; FILING DATE: 20-DEC-1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/JP92/00544
; FILING DATE: 24-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-32084
; FILING DATE: 19-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 3-95476
; FILING DATE: 25-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: WEGNER, Harold C.
; REGISTRATION NUMBER: 25,258
; REFERENCE/DOCKET NUMBER: 53466/126/AAOK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 141:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-137-117D-141

Query Match 24.2%; Score 39; DB 1; Length 32;
Best Local Similarity 50.0%; Pred. No. 48;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```
Qy      3 THTITKLNAENNATFY 18
          | ||: |  |: ||:|
Db     16 TFTISSLQPEDIATYY 31
```

RESULT 8
US-08-436-717-122
; Sequence 122, Application US/08436717
; Patent No. 5817790
; GENERAL INFORMATION:
; APPLICANT: TSUCHIYA, Masayuki
; APPLICANT: SATO, Koh
; APPLICANT: BENDIG, Mary
; APPLICANT: JONES, Steven
; APPLICANT: SALDANHA, Jose
; TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN
; TITLE OF INVENTION: INTERLEUKIN-6 RECEPTOR
; NUMBER OF SEQUENCES: 158
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,717
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/137,117
; FILING DATE: 20-DEC-1993
; APPLICATION NUMBER: WO PCT/JP92/00544
; FILING DATE: 24-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-32084
; FILING DATE: 19-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 3-95476
; FILING DATE: 25-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: WEGNER, Harold C.
; REGISTRATION NUMBER: 25,258
; REFERENCE/DOCKET NUMBER: 53466/126/AAOK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 122:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-436-717-122

Query Match 24.2%; Score 39; DB 2; Length 32;
Best Local Similarity 50.0%; Pred. No. 48;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```
Qy      3 THTITKLNAENNATFY 18
          | ||: |  |: ||:|
Db     16 TFTISSLQPEDIATYY 31
```

RESULT 9
US-08-436-717-124
; Sequence 124, Application US/08436717
; Patent No. 5817790
; GENERAL INFORMATION:
; APPLICANT: TSUCHIYA, Masayuki
; APPLICANT: SATO, Koh
; APPLICANT: BENDIG, Mary
; APPLICANT: JONES, Steven

```
;    APPLICANT:  SALDANHA, Jose
;    TITLE OF INVENTION:  RESHAPED HUMAN ANTIBODY TO HUMAN
;    TITLE OF INVENTION:  INTERLEUKIN-6 RECEPTOR
;    NUMBER OF SEQUENCES:  158
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/436,717
;      FILING DATE:
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/137,117
;      FILING DATE:  20-DEC-1993
;      APPLICATION NUMBER:  WO PCT/JP92/00544
;      FILING DATE:  24-APR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 4-32084
;      FILING DATE:  19-FEB-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 3-95476
;      FILING DATE:  25-APR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  WEGNER, Harold C.
;      REGISTRATION NUMBER:  25,258
;      REFERENCE/DOCKET NUMBER:  53466/126/AAOK
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  124:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  32 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-436-717-124


  Query Match             24.2%;  Score 39;  DB 2;  Length 32;
  Best Local Similarity   50.0%;  Pred. No. 48;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 THTITKLNAENNATFY 18
           | ||: |  |: ||:|
Db      16 TFTISSLQPEDIATYY 31


RESULT  10
US-08-436-717-141
; Sequence 141, Application US/08436717
; Patent No. 5817790
;  GENERAL INFORMATION:
;    APPLICANT:  TSUCHIYA, Masayuki
;    APPLICANT:  SATO, Koh
;    APPLICANT:  BENDIG, Mary
;    APPLICANT:  JONES, Steven
;    APPLICANT:  SALDANHA, Jose
;    TITLE OF INVENTION:  RESHAPED HUMAN ANTIBODY TO HUMAN
;    TITLE OF INVENTION:  INTERLEUKIN-6 RECEPTOR
;    NUMBER OF SEQUENCES:  158
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/436,717
;      FILING DATE:
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/137,117
;      FILING DATE:  20-DEC-1993
;      APPLICATION NUMBER:  WO PCT/JP92/00544
;      FILING DATE:  24-APR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 4-32084
;      FILING DATE:  19-FEB-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 3-95476
;      FILING DATE:  25-APR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  WEGNER, Harold C.
;      REGISTRATION NUMBER:  25,258
;      REFERENCE/DOCKET NUMBER:  53466/126/AAOK
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  141:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  32 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-436-717-141


  Query Match             24.2%;  Score 39;  DB 2;  Length 32;
  Best Local Similarity   50.0%;  Pred. No. 48;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 THTITKLNAENNATFY 18
           | ||: |  |: ||:|
Db      16 TFTISSLQPEDIATYY 31


RESULT  11
US-08-553-501A-67
; Sequence 67, Application US/08553501A
; Patent No. 5856135
;  GENERAL INFORMATION:
;    APPLICANT:  TSUCHIYA, Masayuki
;    APPLICANT:  SATO, Koh
;    APPLICANT:  HIRATA, Yuichi
;    TITLE OF INVENTION:  RESHAPED HUMAN ANTIBODY TO HUMAN
;    TITLE OF INVENTION:  INTERLEUKIN-6
;    NUMBER OF SEQUENCES:  91
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/553,501A
;      FILING DATE:  20-FEB-1996
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/JP94/00859
;      FILING DATE:  30-MAY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 5-129787
;      FILING DATE:  31-MAY-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  WEGNER, Harold C.
;      REGISTRATION NUMBER:  25,258
;      REFERENCE/DOCKET NUMBER:  53466/177/AAOK
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  67:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  32 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-553-501A-67


  Query Match             24.2%;  Score 39;  DB 2;  Length 32;
  Best Local Similarity   50.0%;  Pred. No. 48;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 THTITKLNAENNATFY 18
           | ||: |  |: ||:|
Db      16 TFTISSLQPEDIATYY 31


RESULT  12
US-08-553-501A-71
; Sequence 71, Application US/08553501A
; Patent No. 5856135
;  GENERAL INFORMATION:
;    APPLICANT:  TSUCHIYA, Masayuki
;    APPLICANT:  SATO, Koh
;    APPLICANT:  HIRATA, Yuichi
;    TITLE OF INVENTION:  RESHAPED HUMAN ANTIBODY TO HUMAN
;    TITLE OF INVENTION:  INTERLEUKIN-6
;    NUMBER OF SEQUENCES:  91
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/553,501A
;      FILING DATE:  20-FEB-1996
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/JP94/00859
;      FILING DATE:  30-MAY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 5-129787
;      FILING DATE:  31-MAY-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  WEGNER, Harold C.
;      REGISTRATION NUMBER:  25,258
;      REFERENCE/DOCKET NUMBER:  53466/177/AAOK
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  71:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  32 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-553-501A-71


  Query Match             24.2%;  Score 39;  DB 2;  Length 32;
  Best Local Similarity   50.0%;  Pred. No. 48;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 THTITKLNAENNATFY 18
           | ||: |  |: ||:|
Db      16 TFTISSLQPEDIATYY 31


RESULT  13
US-08-560-558E-3
; Sequence 3, Application US/08560558E
; Patent No. 5891996
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Humanized and chimeric monoclonal
;    TITLE OF INVENTION:  antibodies that recognize epidermal growth factor receptor
;    TITLE OF INVENTION:  EGF-R); diagnostic and therapeutic use.
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Allen C. Turner, TRASK, BRITT & ROSSA
;      STREET:  P.O. Box 2250
;      CITY:  Salt Lake City
;      STATE:  Utah
;      COUNTRY:  United States of America
;      ZIP:  84110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  WINDOWS95
;      SOFTWARE:  WordPerfect 5.1/5.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/560,558E
;      FILING DATE:  No. 5891996ember 17, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Turner, Allen C.
;      REGISTRATION NUMBER:  33,041
;      REFERENCE/DOCKET NUMBER:  2720US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (801) 532-1922
;      TELEFAX:  (801) 531-9168
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  32 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
US-08-560-558E-3


  Query Match             24.2%;  Score 39;  DB 2;  Length 32;
  Best Local Similarity   50.0%;  Pred. No. 48;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 THTITKLNAENNATFY 18
           | ||: |  |: ||:|
Db      16 TFTISSLQPEDIATYY 31
```

```
RESULT 14
US-08-765-783A-89
; Sequence 89, Application US/08765783A
; Patent No. 5994524
;  GENERAL INFORMATION:
;    APPLICANT:  Matsushima, Kouji
;    APPLICANT:  Matsumoto, Yoshihiro
;    APPLICANT:  Yamada, Yoshiki
;    APPLICANT:  Sato, Koh
;    APPLICANT:  Tsuchiya, Masayuki
;    APPLICANT:  Yamazaki, Tatsumi
;    TITLE OF INVENTION:  Reshaped Human Antibody to
;    TITLE OF INVENTION:  Interleukin-8
;    NUMBER OF SEQUENCES:  105
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Avenue, NW, suite 5500
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006-1888
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/765,783A
;      FILING DATE:  07-MAR-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murashige, Kate H
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  35029-20001.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-887-1500
;      TELEFAX:  202-822-0168
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  89:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  32 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-765-783A-89


  Query Match             24.2%;  Score 39;  DB 2;  Length 32;
  Best Local Similarity   50.0%;  Pred. No. 48;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 THTITKLNAENNATFY 18
           | ||: |  |: ||:|
Db      16 TFTISSLQPEDIATYY 31


RESULT 15
US-08-765-783A-91
; Sequence 91, Application US/08765783A
; Patent No. 5994524
;  GENERAL INFORMATION:
;    APPLICANT:  Matsushima, Kouji
;    APPLICANT:  Matsumoto, Yoshihiro
;    APPLICANT:  Yamada, Yoshiki
;    APPLICANT:  Sato, Koh
;    APPLICANT:  Tsuchiya, Masayuki
;    APPLICANT:  Yamazaki, Tatsumi
;    TITLE OF INVENTION:  Reshaped Human Antibody to
;    TITLE OF INVENTION:  Interleukin-8
;    NUMBER OF SEQUENCES:  105
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Avenue, NW, suite 5500
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006-1888
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/765,783A
;      FILING DATE:  07-MAR-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murashige, Kate H
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  35029-20001.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-887-1500
;      TELEFAX:  202-822-0168
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  91:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  32 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-765-783A-91


  Query Match             24.2%;  Score 39;  DB 2;  Length 32;
  Best Local Similarity   50.0%;  Pred. No. 48;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 THTITKLNAENNATFY 18
           | ||: |  |: ||:|
Db      16 TFTISSLQPEDIATYY 31


RESULT 16
US-09-205-231-67
; Sequence 67, Application US/09205231
; Patent No. 6121423
;  GENERAL INFORMATION:
;    APPLICANT:  TSUCHIYA, Masayuki
;    APPLICANT:  SATO, Koh
;    APPLICANT:  HIRATA, Yuichi
;    TITLE OF INVENTION:  RESHAPED HUMAN ANTIBODY TO HUMAN
;    TITLE OF INVENTION:  INTERLEUKIN-6
;    NUMBER OF SEQUENCES:  91
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/205,231
;      FILING DATE:
;      CLASSIFICATION:
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/553,501
;      FILING DATE:  20-FEB-1996
;      APPLICATION NUMBER:  WO PCT/JP94/00859
;      FILING DATE:  30-MAY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 5-129787
;      FILING DATE:  31-MAY-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  WEGNER, Harold C.
;      REGISTRATION NUMBER:  25,258
;      REFERENCE/DOCKET NUMBER:  53466/177/AAOK
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  67:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  32 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-205-231-67


  Query Match             24.2%;  Score 39;  DB 3;  Length 32;
  Best Local Similarity   50.0%;  Pred. No. 48;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 THTITKLNAENNATFY 18
           | ||: |  |: ||:|
Db      16 TFTISSLQPEDIATYY 31


RESULT  17
US-09-205-231-71
; Sequence 71, Application US/09205231
; Patent No. 6121423
;  GENERAL INFORMATION:
;    APPLICANT:  TSUCHIYA, Masayuki
;    APPLICANT:  SATO, Koh
;    APPLICANT:  HIRATA, Yuichi
;    TITLE OF INVENTION:  RESHAPED HUMAN ANTIBODY TO HUMAN
;    TITLE OF INVENTION:  INTERLEUKIN-6
;    NUMBER OF SEQUENCES:  91
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/205,231
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/553,501
;      FILING DATE:  20-FEB-1996
;      APPLICATION NUMBER:  WO PCT/JP94/00859
;      FILING DATE:  30-MAY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 5-129787
;      FILING DATE:  31-MAY-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  WEGNER, Harold C.
;      REGISTRATION NUMBER:  25,258
;      REFERENCE/DOCKET NUMBER:  53466/177/AAOK
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  71:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  32 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-205-231-71


  Query Match             24.2%;  Score 39;  DB 3;  Length 32;
  Best Local Similarity   50.0%;  Pred. No. 48;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 THTITKLNAENNATFY 18
           | ||: |  |: ||:|
Db      16 TFTISSLQPEDIATYY 31


RESULT  18
US-08-646-265A-123
; Sequence 123, Application US/08646265A
; Patent No. 6214973
;  GENERAL INFORMATION:
;    APPLICANT:  OHTOMO, Toshihiko
;    APPLICANT:  SATO, Koh
;    APPLICANT:  TSUCHIYA, Masayuki
;    TITLE OF INVENTION:  RESHAPED HUMAN ANTIBODY TO HUMAN
;    TITLE OF INVENTION:  MEDULLOBLASTOMA CELLS
;    NUMBER OF SEQUENCES:  132
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/646,265A
;      FILING DATE:  09-SEP-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/JP94/01763
;      FILING DATE:  19-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 5-291078
;      FILING DATE:  19-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  WEGNER, Harold C.
;      REGISTRATION NUMBER:  25,258
;      REFERENCE/DOCKET NUMBER:  53466/184
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  123:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  32 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-646-265A-123


  Query Match             24.2%;  Score 39;  DB 4;  Length 32;
```

```
  Best Local Similarity   50.0%;  Pred. No. 48;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 THTITKLNAENNATFY 18
           | ||: |  |: ||:|
Db      16 TFTISSLQPEDIATYY 31


RESULT  19
US-09-416-557-89
; Sequence 89, Application US/09416557
; Patent No. 6245894
;  GENERAL INFORMATION:
;    APPLICANT:  Matsushima, Kouji
;    APPLICANT:  Matsumoto, Yoshihiro
;    APPLICANT:  Yamada, Yoshiki
;    APPLICANT:  Sato, Koh
;    APPLICANT:  Tsuchiya, Masayuki
;    APPLICANT:  Yamazaki, Tatsumi
;    TITLE OF INVENTION:  Reshaped Human Antibody to
;    TITLE OF INVENTION:  Interleukin-8
;    NUMBER OF SEQUENCES:  105
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Avenue, NW, suite 5500
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006-1888
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/416,557
;      FILING DATE:  12-October-1999
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/765,783
;      FILING DATE:  7-March-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murashige, Kate H
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  35029-20001.10
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-887-1500
;      TELEFAX:  202-822-0168
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  89:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  32 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-416-557-89


  Query Match             24.2%;  Score 39;  DB 4;  Length 32;
  Best Local Similarity   50.0%;  Pred. No. 48;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 THTITKLNAENNATFY 18
           | ||: |  |: ||:|
Db      16 TFTISSLQPEDIATYY 31


RESULT  20
US-09-416-557-91
; Sequence 91, Application US/09416557
; Patent No. 6245894
;  GENERAL INFORMATION:
;    APPLICANT:  Matsushima, Kouji
;    APPLICANT:  Matsumoto, Yoshihiro
;    APPLICANT:  Yamada, Yoshiki
;    APPLICANT:  Sato, Koh
;    APPLICANT:  Tsuchiya, Masayuki
;    APPLICANT:  Yamazaki, Tatsumi
;    TITLE OF INVENTION:  Reshaped Human Antibody to
;    TITLE OF INVENTION:  Interleukin-8
;    NUMBER OF SEQUENCES:  105
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Avenue, NW, suite 5500
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006-1888
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/416,557
;      FILING DATE:  12-October-1999
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/765,783
;      FILING DATE:  7-March-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murashige, Kate H
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  35029-20001.10
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-887-1500
;      TELEFAX:  202-822-0168
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  91:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  32 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-416-557-91


  Query Match             24.2%;  Score 39;  DB 4;  Length 32;
  Best Local Similarity   50.0%;  Pred. No. 48;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 THTITKLNAENNATFY 18
           | ||: |  |: ||:|
Db      16 TFTISSLQPEDIATYY 31


Search completed: July  8, 2002, 11:49:17
Job time: 313 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 8, 2002, 11:39:08 ; Search time 48.86 Seconds
(without alignments)
27.280 Million cell updates/sec

Title: US-09-461-061A-2
Perfect score: 60
Sequence: 1 TLTHTITKLNAE 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 158732

Minimum DB seq length: 0
Maximum DB seq length: 12

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	60	100.0	12	21	AAY95406	Anti-angiogenic pe
2	30	50.0	10	22	AAJ01502	Hepatitis C virus
3	27	45.0	10	22	AAG96132	Human complementar
4	27	45.0	10	22	AAG84421	Arabidopsis thalia
5	26	43.3	9	16	AAR61413	TNF-alpha related
6	26	43.3	9	19	AAW58653	Platelet factor 4
7	26	43.3	11	16	AAR87682	h-TNF (70-80). Sy
8	26	43.3	11	16	AAR87684	h-TNF (70-80) anal
9	26	43.3	11	21	AAY52581	Amaranthus viridis
10	26	43.3	12	18	AAW35497	TNF alpha peptide
11	25	41.7	9	20	AAY46762	Immunogenic peptid
12	24	40.0	8	22	AAG89420	p53 DR supermotif
13	24	40.0	8	22	AAG89421	p53 DR supermotif
14	24	40.0	10	18	AAW30080	Hamster neurturin
15	24	40.0	10	18	AAW13715	Internal neurturin
16	24	40.0	10	20	AAY16636	WO9914235 Seq ID N
17	24	40.0	10	22	AAM42846	Mycoplasma genital
18	24	40.0	11	16	AAR87683	h-TNF (70-80) anal
19	24	40.0	11	18	AAW35449	Cytokine derived p
20	24	40.0	11	19	AAW47911	Human tumour necro
21	24	40.0	11	19	AAW47912	Human tumour necro
22	24	40.0	11	19	AAW47916	Human tumour necro
23	24	40.0	11	19	AAW47917	Human tumour necro
24	24	40.0	11	19	AAW47918	Human tumour necro
25	24	40.0	11	19	AAW47922	Human tumour necro
26	24	40.0	11	19	AAW47900	Human tumour necro
27	24	40.0	11	19	AAW47901	Human tumour necro
28	24	40.0	11	19	AAW47903	Human tumour necro
29	24	40.0	11	19	AAW47904	Human tumour necro
30	24	40.0	11	19	AAW47876	Human tumour necro
31	24	40.0	11	19	AAW47877	Human tumour necro
32	24	40.0	11	19	AAW47860	Human tumour necro
33	24	40.0	11	19	AAW47865	Human tumour necro
34	24	40.0	11	19	AAW47863	Human tumour necro
35	24	40.0	11	19	AAW47868	Human tumour necro
36	24	40.0	11	19	AAW47869	Human tumour necro
37	24	40.0	11	19	AAW47870	Human tumour necro
38	24	40.0	11	19	AAW47871	Human tumour necro
39	24	40.0	11	19	AAW47872	Human tumour necro
40	24	40.0	11	19	AAW47873	Human tumour necro
41	24	40.0	11	19	AAW47874	Human tumour necro
42	24	40.0	11	19	AAW47875	Human tumour necro
43	24	40.0	11	22	AAB84078	Immune modulating
44	24	40.0	12	19	AAW47910	Human tumour necro
45	24	40.0	12	19	AAW47892	Human tumour necro

ALIGNMENTS

RESULT 1
AAY95406
ID AAY95406 standard; Peptide; 12 AA.
XX
AC AAY95406;
XX
DT 25-SEP-2000 (first entry)
XX
DE Anti-angiogenic peptide N-terminal fragment.
XX
KW Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;
KW endothelial cell proliferation; apoptosis; cancer; ocular disorder;
KW rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic;
KW therapy; human; D3 peptide.
XX
OS Homo sapiens.
XX
PN WO200035407-A2.
XX
PD 22-JUN-2000.
XX
PF 02-DEC-1999; 99WO-US28465.
XX
PR 16-DEC-1998; 98US-0112427.
XX
PA (UTEM) UNIV TEMPLE.
PA (MCCR/) MCCRAE R K.
XX
PI McCrae RK;
XX
DR WPI; 2000-442247/38.
XX
PT Composition for inhibiting angiogenesis and endothelial cell
PT proliferation, inducing endothelial cell apoptosis and treating cancer,

PT rheumatoid arthritis, and ocular disorders comprises a kininogen domain
PT 3 analog -
XX
PS Claim 3; Page 25; 44pp; English.
XX
CC The present sequence is that of an N-terminal fragment of a novel
CC anti-angiogenic D3 peptide (see AAY95408) derived from human high
CC mol.wt. kininogen (HK) domain 3 (see AAY95426). The full-length D3
CC peptide inhibits endothelial cell proliferation and thus possesses
CC anti-angiogenic activity. It is an example of peptides of the
CC invention (see AAY95405-26) that are analogues of certain sites in
CC the HK domain 3. The peptides inhibit endothelial cell proliferation
CC and may also induce endothelial cell apoptosis. Compositions
CC including the peptides are used in claimed methods for inhibiting
CC angiogenesis, inhibiting endothelial cell proliferation, and
CC inducing endothelial cell apoptosis. Cancer, rheumatoid arthritis,
CC and ocular disorders characterized by undesired vascularization of
CC the retina are treated.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 60; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLTHTITKLNAE 12
||||||||||||
Db 1 tlthtitklnae 12

RESULT 2
AAJ01502
ID AAJ01502 standard; Peptide; 10 AA.
XX
AC AAJ01502;
XX
DT 02-JUL-2001 (first entry)
XX
DE Hepatitis C virus epitope #1493.
XX
KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX
OS Hepatitis C virus.
XX
PN WO200121189-A1.
XX
PD 29-MAR-2001.
XX
PF 19-JUL-2000; 2000WO-US19774.
XX
PR 19-JUL-1999; 99US-0357737.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-308046/32.
XX
PT A new composition useful as a vaccines against hepatitis C virus -
XX
PS Disclosure; Page 138; 214pp; English.
XX
CC The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention.
XX
SQ Sequence 10 AA;

Query Match 50.0%; Score 30; DB 22; Length 10;
Best Local Similarity 85.7%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LTHTITK 8
||| |||
Db 1 lthpitk 7

RESULT 3
AAG96132
ID AAG96132 standard; Peptide; 10 AA.
XX
AC AAG96132;
XX
DT 18-SEP-2001 (first entry)
XX
DE Human complementary peptide, SEQ ID NO: 2326.
XX
KW Human; complementary peptide; ligand; drug discovery; drug design.
XX
OS Homo sapiens.
XX
PN WO200142277-A2.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-GB04776.
XX
PR 13-DEC-1999; 99GB-0029464.
XX
PA (PROT-) PROTEOM LTD.
XX
PI Roberts GW, Heal JR;
XX
DR WPI; 2001-408419/43.
XX
PT A set of peptide ligands consisting of specific complementary peptides
PT to proteins encoded by genes of the human genome, useful in an assay
PT for screening and identifying of one or more novel peptides which are
PT drug candidates or pro-drugs -
XX
PS Example 4; Page 375; 646pp; English.
XX
CC The invention relates to a set of complementary peptide ligands
CC generated from the human genome. The complementary peptides
CC interact with their relevant target proteins encoded in the human
CC genome. They can be used as reagents in drug discovery and as lead
CC ligands to facilitate drug design and development. The present
CC sequence is a complementary peptide provided in the specification.
XX
SQ Sequence 10 AA;

Query Match 45.0%; Score 27; DB 22; Length 10;
Best Local Similarity 44.4%; Pred. No. 1.5e+02;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TLTHTITKL 9
| ||:::|
Db 2 tathslsrl 10

RESULT 4
AAG84421
ID AAG84421 standard; Peptide; 10 AA.
XX
AC AAG84421;
XX
DT 11-SEP-2001 (first entry)
XX
DE Arabidopsis thaliana peptide ligand #1061.

XX
KW Plant; peptide pesticide; peptide herbicide; agricultural research.
XX
OS Arabidopsis thaliana.
XX
PN WO200142279-A2.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-GB04781.
XX
PR 13-DEC-1999; 99GB-0029469.
XX
PA (PROT-) PROTEOM LTD.
XX
PI Roberts GW, Heal JR;
XX
DR WPI; 2001-381629/40.
XX
PT A set of peptide ligands for agricultural research and development of
PT therapeutic agents comprise specific complementary peptides to proteins
PT encoded by genes of plant genomes -
XX
PS Example 4; Page 179; 201pp; English.
XX
CC The present invention relates to a set of peptide ligands consisting of
CC specific complementary peptides to proteins encoded by genes of plant
CC genomes. The present sequence is one such peptide from Arabidopsis
CC thaliana. The peptides of the present invention are useful in an assay to
CC identify a peptide, especially a peptide pesticide or herbicide. The
CC peptides are also useful for tools for agricultural research and
CC development.
XX
SQ Sequence 10 AA;

Query Match 45.0%; Score 27; DB 22; Length 10;
Best Local Similarity 62.5%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TLTHTITK 8
|| ||: |
Db 3 tlqhtlrk 10

RESULT 5
AAR61413
ID AAR61413 standard; peptide; 9 AA.
XX
AC AAR61413;
XX
DT 25-AUG-1995 (first entry)
XX
DE TNF-alpha related peptide designated CT-119.
XX
KW antiinflammatory; platelet factor 4; PF-4 autoimmune disease;
KW graft-versus-host; reperfusion injury; atherosclerosis; asthma.
XX
OS Synthetic.
XX
PN WO9500543-A.
XX
PD 05-JAN-1995.
XX
PF 17-JUN-1994; 94WO-US06888.
XX
PR 18-JUN-1993; 93US-0080371.
XX
PA (COUN/) COUNTS D F.
PA (DUFF/) DUFF R G.
XX
PI Counts DF, Duff RG;
XX
DR WPI; 1995-052005/07.
XX
PT New peptide(s) and derivs. based on platelet factor 4 - used for
PT inhibiting an immune response, including an inflammatory response
PT in e.g. auto:immune diseases.
XX
PS Disclosure; Page 38; 99pp; English.
XX
CC New peptides are disclosed which include any peptide, peptide
CC derivative or peptide analogue which comprises either (i) at least a
CC 4 amino acid portion of PF-4 (see AAR61401) or a functionally equivalent
CC sequence, or (ii) at least a 6 amino acid sequence which is at least
CC 66% homologous to a portion of the PF-4 sequence, or a functionally
CC equivalent sequence. Pref. the peptide contains the sequence
CC Thr-Ser-Gln and/or Val-Arg-Pro, and more preferably Thr-Thr-Ser-Gln
CC and/or Val-Arg-Pro-Arg. The most preferred peptide is
CC Thr-Thr-Ser-Gln-Val-Arg-Pro-Arg (AAR61393), designated CT-112.
CC The peptide may be derivatised at the N- and/or C-terminal, or may be
CC cyclised, substituted, truncated or contain D-amino acid residues.
CC The peptides exhibit antiinflammatory activity and may be used to treat
CC autoimmune diseases (such as insulin-dependent diabetes, ulcerative
CC colitis, rheumatoid arthritis, scleroderma, mixed connective tissue
CC disease and SLE), reperfusion tissue damage, inflammatory lung disease,
CC graft-versus-host disease, atherosclerosis and asthma.
CC The present sequence, designated CT-119, is a comparison sequence
CC from TNF-alpha, constructed for testing and comparison purposes.
CC When tested for antiinflammatory activity by the mouse ear acute
CC inflammation model at a dose of 60 mg/kg administered gradually
CC over 24 hr, the peptide CT-112 gave an inhibition of 78%. In
CC comparison the present sequence gave a value of 74%.
XX
SQ Sequence 9 AA;

Query Match 43.3%; Score 26; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LTHTI 6
|||||
Db 5 lthti 9

RESULT 6
AAW58653
ID AAW58653 standard; peptide; 9 AA.
XX
AC AAW58653;
XX
DT 08-SEP-1998 (first entry)
XX
DE Platelet factor 4 derived peptide SEQ ID NO:21 CT-119.
XX
KW Platelet factor 4; PF4; anti-inflammatory; inhibition; inflammation;
KW autoimmune disease; graft versus host disease; reperfusion injury;
KW atherosclerosis; asthma; chemokine.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN US5776892-A.
XX
PD 07-JUL-1998.
XX
PF 16-JUN-1994; 94US-0250550.
XX
PR 16-JUN-1994; 94US-0259550.
PR 21-DEC-1990; 90US-0631823.
PR 24-MAR-1993; 93US-0037486.
PR 18-JUN-1993; 93US-0080371.
XX
PA (CURA-) CURATIVE HEALTH SERVICES INC.
XX

```
PI   Counts DF,  Duff RG;
XX
DR   WPI; 1998-398086/34.
XX
PT   New antiinflammatory peptide(s) based on platelet factor 4 sequences
PT   - used for treating e.g. auto:immune diseases, graft versus host
PT   disease, reperfusion injury, atherosclerosis or asthma
XX
PS   Example 14; Column 21; 55pp; English.
XX
CC   The present sequence represents a peptide which is related to platelet
CC   factor 4 (PF4) and can be used to inhibit an inflammatory response.
CC   Peptides, peptide analogues and peptide derivatives of PF4 can be used
CC   for treating e.g. autoimmune diseases such as insulin dependent
CC   diabetes, Goodpasture's syndrome, pemphigus and pemphigoid, primary
CC   biliary cirrhosis, ulcerative colitis, rheumatoid arthritis,
CC   scleroderma, mixed connective tissue disease and lupus erythematosus,
CC   graft versus host disease, septic shock, reperfusion injury (including
CC   injury subsequent to myocardial or cerebral infarction),
CC   atherosclerosis, asthma and inflammatory lung disease. The peptides
CC   give a new and effective method of inhibiting the inflammatory response
CC   by acting on cytokines rather than the prior art arachidonic acid.
XX
SQ   Sequence   9 AA;


  Query Match             43.3%;  Score 26;  DB 19;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 6.4e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 LTHTI 6
          |||||
Db      5 lthti 9


RESULT   7
AAR87682
ID   AAR87682 standard; peptide; 11 AA.
XX
AC   AAR87682;
XX
DT   10-MAY-1996  (first entry)
XX
DE   h-TNF (70-80).
XX
KW   TNF; tumour necrosis factor; neutrophil; superoxide; inflammation;
KW   AIDS, cancer; rheumatoid arthritis; ARDS; cystic fibrosis;
KW   tuberculosis; TB; Crohn's disease; sarcoidosis; leprosy; syphilis;
KW   malaria; adult respiratory distress syndrome.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "this residue is optionally N-acylated or
FT                          glycosylated"
FT   Modified-site   11
FT                   /note= "this residue is optionally in amide or
FT                          ester form and may be glycosylated"
XX
PN   AU9344664-A.
XX
PD   02-MAR-1995.
XX
PF   12-MAR-1991;   91AU-0074762.
XX
PR   12-MAR-1991;   91AU-0074762.
PR   13-AUG-1993;   93AU-0044664.
XX
PA   (PEPT-) PEPTIDE TECHNOLOGY LTD.
XX
PI   Ferrante A,  Rathjen DA,  Widmer F;
XX
DR   WPI; 1995-115742/16.
XX
PT   Novel peptides which stimulate neutrophils - are fragments of
PT   tumour necrosis factor
XX
PS   Claim 3; Page 32; 51pp; English.
XX
CC   New peptides are provided which prime neutrophils for superoxide
CC   production and an enhanced respiratory burst following treatment
CC   with N-formyl-L-methionine-L-leucyl-L-phenylalanine. The peptides are
CC   described generically. They are fragments from the region of amino acids
CC   54 to 94 of human TNF, or analogues of such fragments. Their
CC   neutrophil-stimulating activity can be used in the treatment of diseases
CC   such as AIDS, cancer, rheumatoid arthritis, ARDS, cystic fibrosis,
CC   tuberculosis, Crohn's disease, sarcoidosis, leprosy, syphilis and
CC   malaria. The present sequence is one of 6 specific peptides having the
CC   above activity and which fit the the generic formula.
XX
SQ   Sequence   11 AA;


  Query Match             43.3%;  Score 26;  DB 16;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 LTHTI 6
          |||||
Db      7 lthti 11


RESULT   8
AAR87684
ID   AAR87684 standard; peptide; 11 AA.
XX
AC   AAR87684;
XX
DT   10-MAY-1996  (first entry)
XX
DE   h-TNF (70-80) analogue.
XX
KW   TNF; tumour necrosis factor; neutrophil; superoxide; inflammation;
KW   AIDS, cancer; rheumatoid arthritis; ARDS; cystic fibrosis;
KW   tuberculosis; TB; Crohn's disease; sarcoidosis; leprosy; syphilis;
KW   malaria; adult respiratory distress syndrome.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "this residue is optionally N-acylated or
FT                          glycosylated"
FT   Modified-site   11
FT                   /note= "this residue is optionally in amide or
FT                          ester form and may be glycosylated"
XX
PN   AU9344664-A.
XX
PD   02-MAR-1995.
XX
PF   12-MAR-1991;   91AU-0074762.
XX
PR   12-MAR-1991;   91AU-0074762.
PR   13-AUG-1993;   93AU-0044664.
XX
PA   (PEPT-) PEPTIDE TECHNOLOGY LTD.
XX
PI   Ferrante A,  Rathjen DA,  Widmer F;
XX
DR   WPI; 1995-115742/16.
XX
PT   Novel peptides which stimulate neutrophils - are fragments of
PT   tumour necrosis factor
XX
```

PS Claim 5; Page 32; 51pp; English.
XX
CC New peptides are provided which prime neutrophils for superoxide
CC production and an enhanced respiratory burst following treatment
CC with N-formyl-L-methionine-L-leucyl-L-phenylalanine. The peptides are
CC described generically. They are fragments from the region of amino acids
CC 54 to 94 of human TNF, or analogues of such fragments. Their
CC neutrophil-stimulating activity can be used in the treatment of diseases
CC such as AIDS, cancer, rheumatoid arthritis, ARDS, cystic fibrosis,
CC tuberculosis, Crohn's disease, sarcoidosis, leprosy, syphilis and
CC malaria. The present sequence is one of 6 specific peptides having the
CC above activity and which fit the the generic formula.
XX
SQ Sequence 11 AA;

Query Match 43.3%; Score 26; DB 16; Length 11;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LTHTI 6
|||||
Db 7 lthti 11

RESULT 9
AAY52581
ID AAY52581 standard; peptide; 11 AA.
XX
AC AAY52581;
XX
DT 28-FEB-2000 (first entry)
XX
DE Amaranthus viridis amarandin-1 N-terminal 11-mer peptide.
XX
KW Amarandin; N-glycosidase; ribosome; inactivation; antiviral;
KW inactivation; protein synthesis; N-glycosidase; adenine; ribosomal RNA;
KW elongation factor 2; EF2-dependent GTPase; irreversible; inhibition;
KW infectivity; reduction; replication; plant virus; tobacco mosaic virus;
KW watermelon mosaic virus; cauliflower mosaic virus; potato virus X;
KW potato virus Y; potato leafroll virus; cucumber mosaic virus;
KW human disease; HIV-1; recombinant expression.
XX
OS Synthetic.
OS Amaranthus viridis.
XX
PN US6001986-A.
XX
PD 14-DEC-1999.
XX
PF 22-AUG-1997; 97US-0916443.
XX
PR 22-AUG-1997; 97US-0916443.
XX
PA (KOKU-) KOREA KUMHO PETROCHEMICAL CO LTD.
XX
PI Kim YS, Cho JW, Chung CH, Park SC, Oh SK, Lee H;
XX
DR WPI; 2000-061906/05.
XX
PT Isolated DNA encoding antiviral proteins amarandin 1 and 2, from
PT Amaranthus viridis -
XX
PS Claim 5; Column 3; 19pp; English.
XX
CC This sequence represents the N-terminal 11 residues of amarandin-1, an
CC antiviral protein from Amarathus viridis. Amarandin-1 (AAY52586) and
CC amarandin-2 (AAY52587) are inactivators of eukaryotic ribosomes. They
CC have N-glycosidase activity, cleaving the N-glycosidic bond of
CC adenine in a specific ribosomal RNA sequence, thus modifying
CC the EF (elongation factor)-2-dependent GTPase activity of
CC the ribosomal subunit. This irreversibly impairs protein
CC synthesis. Amarandins 1 and 2 have been shown to reduce infectivity of
CC tobacco mosaic virus and other plant viruses including watermelon
CC mosaic virus, cauliflower mosaic virus, potato virus X and Y, potato
CC leafroll virus and cucumber mosaic virus. These proteins may also be
CC very important to inhibit replication by human viruses, e.g., HIV-1.
CC DNA encoding amarandins 1 and 2 may be used in expression vectors for
CC the recombinant expression of these antiviral proteins in a host cell.
XX
SQ Sequence 11 AA;

Query Match 43.3%; Score 26; DB 21; Length 11;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LTHTITK 8
|| |:||
Db 3 ltftvtk 9

RESULT 10
AAW35497
ID AAW35497 standard; peptide; 12 AA.
XX
AC AAW35497;
XX
DT 22-APR-1998 (first entry)
XX
DE TNF alpha peptide from WO9738011.
XX
KW T-cell stimulatory peptide; immunogen; non-dendritic; carrier; tumour;
KW scaffold; inhibition; metastasis; wound healing; solid phase.
XX
OS Unidentified.
XX
PN WO9738011-A1.
XX
PD 16-OCT-1997.
XX
PF 03-APR-1997; 97WO-DE00146.
XX
PR 03-APR-1996; 96DK-0000398.
XX
PA (PEPR-) PEPRESEARCH AS.
XX
PI Heegaard PMH, Jakobsen PH;
XX
DR WPI; 1997-512645/47.
XX
PT Non-dendritic peptide carrier linked to a solid phase - useful as a
PT diagnostic agent and as a scaffold for production of chemical
PT derivatives
XX
PS Example 7; Page 99; 262pp; English.
XX
CC A non-dendritic peptide carrier (A) has been developed which is coupled
CC through a linker to a solid phase, forming a complex of (A)-solid phase.
CC Where (A) comprises 10-50 amino acids capable of forming a secondary
CC structure in a benign buffer after liberation from the solid phase, and
CC further the (A)-solid phase complex comprises an immunogenic substance
CC and/or an immune mediator coupled on (A). The present sequence
CC represents a peptide used in an example from the present invention. An
CC (A)-solid phase complex can be used as a scaffold for the production of
CC chemical derivatives, characterised by covalently attaching molecules at
CC attachment points. Alternatively (A) is used as a scaffold-peptide for
CC the incorporation into an Immunostimulating Complex (Iscom) resulting an
CC (A)-Iscom complex which is used for the chemical coupling of antigenic
CC substances in an aqueous solution by conjugation. (A) derivatised with
CC one or more peptides having fibronectin-, laminin- or vitronectin-like
CC binding activities can be used for the promotion of cell-attachment to
CC plastic surfaces, in particular to inhibit tumour growth and metastasis,
CC and for promotion of wound healing. Also a derivatised (A) can be used
CC for the selection of specifically-binding aptamers or as a diagnostic
CC agent. Such diagnostic-(A) molecules could be used to detect molecules

```
CC   derived from or indicative of pregnancy or of a disease, such as an
CC   infectious, autoimmune or cancerous disease.
XX
SQ   Sequence   12 AA;


  Query Match             43.3%;  Score 26;  DB 18;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 LTHTI 6
          |||||
Db      8 lthti 12


RESULT  11
AAY46762
ID   AAY46762 standard; Peptide; 9 AA.
XX
AC   AAY46762;
XX
DT   01-DEC-1999  (first entry)
XX
DE   Immunogenic peptide having a human leukocyte antigen binding motif #1373.
XX
KW   Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
KW   immune response; T cell activation; major histocompatibility complex;
KW   cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
KW   prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
KW   vaccine; immunisation.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9945954-A1.
XX
PD   16-SEP-1999.
XX
PF   13-MAR-1998;   98WO-US05039.
XX
PR   13-MAR-1998;   98WO-US05039.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Sette A,  Kubo RT,  Sidney J,  Celis E,  Grey HM,  Southwood S;
XX
DR   WPI; 1999-551214/46.
XX
PT   New immunogenic peptides with HLA binding motif, useful in treatment
PT   and diagnosis of cancers and viral diseases  -
XX
PS   Claim 1; Page 84; 150pp; English.
XX
CC   AAY45390 to AAY48214 represent specifically claimed immunogenic peptides
CC   having a human major histocompatibility complex (MHC) Class I (also
CC   known as human leukocyte antigen (HLA)) binding motif. The immunogenic
CC   peptides can bind to a specific HLA allele (i.e. HLA-A subtypes
CC   HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell
CC   response against the antigen from which the peptide is derived.
CC   Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are
CC   normally induced by an antigen in the form of a peptide fragment bound
CC   to a HLA molecule, rather than the intact foreign antigen itself, and
CC   are particularly important in tumour rejection and in fighting viral
CC   infections. The peptides are therefore useful therapeutically to treat
CC   or prevent viral infections and cancers in mammals (especially humans)
CC   e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
CC   They can be administered as vaccines to elicit an immune response in
CC   individuals susceptible or otherwise at risk of viral infection or
CC   cancer, or used to treat chronic or acute conditions. They are also
CC   useful diagnostically, and can be used to induce a cytotoxic T cell
CC   response, by contacting a cytotoxic T cell with the peptide e.g. to
CC   produce CTLs ex vivo for infusion back into a patient. The
CC   polynucleotides encoding the immunogenic peptides are also useful
CC   therapeutically and for immunisation as above.
XX
SQ   Sequence   9 AA;


  Query Match             41.7%;  Score 25;  DB 20;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 6.4e+05;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TLTHTITK 8
          :| | :||
Db      2 slvhnltk 9


RESULT  12
AAG89420
ID   AAG89420 standard; Peptide; 8 AA.
XX
AC   AAG89420;
XX
DT   11-SEP-2001  (first entry)
XX
DE   p53 DR supermotif binding peptide core sequence #36.
XX
KW   Human; p53; tumour suppressor; human leukocyte antigen; HLA; CTL;
KW   cytotoxic T lymphocyte; T cell; immune response; tumour; cancer;
KW   vaccine; epitope; cytostatic.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200141788-A1.
XX
PD   14-JUN-2001.
XX
PF   11-DEC-2000; 2000WO-US33629.
XX
PR   10-DEC-1999;   99US-0458297.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Fikes J,  Sette A,  Sidney J,  Southwood S,  Chesnut R,  Celis E;
PI   Keogh E;
XX
DR   WPI; 2001-381493/40.
XX
PT   Epitope-based vaccines comprising P53 epitope having a specified
PT   sequences, useful for treating and preventing cancer, the epitopic
PT   peptides is useful as diagnostic agents and for evaluating immune
PT   response  -
XX
PS   Disclosure; Page 111; 138pp; English.
XX
CC   The present invention describes isolated prepared P53 epitopes (I). Also
CC   described are: (1) a cytotoxic T lymphocyte (CTL) that is cultured
CC   in vivo and binds to a complex of (I); (2) a peptide (II) comprising (I)
CC   and a second epitope and has less than 50 contiguous amino acids; (3) a
CC   vaccine composition comprising (II), a unit dose of a peptide with less
CC   than 50 contiguous amino acids with 100% identity to the native peptide
CC   sequence of p53, and a pharmaceutical excipient; (4) an isolated nucleic
CC   acid encoding (I); and (5) an isolated nucleic acid encoding (II). (I)
CC   has cytostatic activity and can be used in vaccines. The vaccine
CC   composition is useful for treating or preventing cancer. (I) and (II)
CC   are useful as diagnostic agents and for evaluating immune responses.
CC   Unlike conventional epitopes, immunosuppressive epitopes that may be
CC   present in whole antigens can be avoided with the use of the vaccine
CC   composition of (I). The ability to combine selected epitopes and
CC   further, to modify the composition of the epitopes enhances the
CC   immunogenicity. The possible pathological side effects caused by
CC   infectious agents or whole protein antigens, which might have their own
CC   intrinsic biological activity, are eliminated. AAG89363 to AAG89747
CC   represent amino acid sequences used in the exemplification of the
CC   present invention.
```

```
XX
SQ   Sequence   8 AA;

  Query Match             40.0%;  Score 24;  DB 22;  Length 8;
  Best Local Similarity   80.0%;  Pred. No. 6.4e+05;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 LTHTI 6
          ||||:
Db      1 lthtl 5


RESULT  13
AAG89421
ID   AAG89421 standard; Peptide; 8 AA.
XX
AC   AAG89421;
XX
DT   11-SEP-2001  (first entry)
XX
DE   p53 DR supermotif binding peptide core sequence #37.
XX
KW   Human; p53; tumour suppressor; human leukocyte antigen; HLA; CTL;
KW   cytotoxic T lymphocyte; T cell; immune response; tumour; cancer;
KW   vaccine; epitope; cytostatic.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200141788-A1.
XX
PD   14-JUN-2001.
XX
PF   11-DEC-2000; 2000WO-US33629.
XX
PR   10-DEC-1999;   99US-0458297.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Fikes J,  Sette A,  Sidney J,  Southwood S,  Chesnut R,  Celis E;
PI   Keogh E;
XX
DR   WPI; 2001-381493/40.
XX
PT   Epitope-based vaccines comprising P53 epitope having a specified
PT   sequences, useful for treating and preventing cancer, the epitopic
PT   peptides is useful as diagnostic agents and for evaluating immune
PT   response  -
XX
PS   Disclosure; Page 111; 138pp; English.
XX
CC   The present invention describes isolated prepared P53 epitopes (I). Also
CC   described are: (1) a cytotoxic T lymphocyte (CTL) that is cultured
CC   in vivo and binds to a complex of (I); (2) a peptide (II) comprising (I)
CC   and a second epitope and has less than 50 contiguous amino acids; (3) a
CC   vaccine composition comprising (II), a unit dose of a peptide with less
CC   than 50 contiguous amino acids with 100% identity to the native peptide
CC   sequence of p53, and a pharmaceutical excipient; (4) an isolated nucleic
CC   acid encoding (I); and (5) an isolated nucleic acid encoding (II). (I)
CC   has cytostatic activity and can be used in vaccines. The vaccine
CC   composition is useful for treating or preventing cancer. (I) and (II)
CC   are useful as diagnostic agents and for evaluating immune responses.
CC   Unlike conventional epitopes, immunosuppressive epitopes that may be
CC   present in whole antigens can be avoided with the use of the vaccine
CC   composition of (I). The ability to combine selected epitopes and
CC   further, to modify the composition of the epitopes enhances the
CC   immunogenicity. The possible pathological side effects caused by
CC   infectious agents or whole protein antigens, which might have their own
CC   intrinsic biological activity, are eliminated. AAG89363 to AAG89747
CC   represent amino acid sequences used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   8 AA;

  Query Match             40.0%;  Score 24;  DB 22;  Length 8;
  Best Local Similarity   80.0%;  Pred. No. 6.4e+05;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 LTHTI 6
          ||||:
Db      2 lthtl 6


RESULT  14
AAW30080
ID   AAW30080 standard; Peptide; 10 AA.
XX
AC   AAW30080;
XX
DT   27-MAR-1998  (first entry)
XX
DE   Hamster neurturin internal peptide P3.
XX
KW   Neurturin; glial-derived neurotrophic factor; GDNF; persephin;
KW   hamster.
XX
OS   Cricetulus sp.
XX
PN   WO9733911-A1.
XX
PD   18-SEP-1997.
XX
PF   14-MAR-1997;   97WO-US03461.
XX
PR   14-MAR-1996;   96US-0615944.
XX
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Johnson EM,  Kotzbauer PT,  Lampe PA,  Milbrandt JD;
XX
DR   WPI; 1997-470818/43.
XX
PT   GDNF-neurturin family related growth factor, Persephin - used to
PT   prevent or treat cellular, neuronal or non-neuronal, degeneration or
PT   insufficiency
XX
PS   Example 4; Page 125; 228pp; English.
XX
CC   This peptide sequence comprises internal peptide P3 of hamster
CC   neurturin isolated from CHO cell conditioned medium.  Other
CC   internal peptides (see AAW30078-79) and an N-terminal peptide (see
CC   AAW30077) of hamster neurturin have also been isolated. Probes and
CC   primers based on these peptide sequences and used to obtain
CC   neurturin cDNA clones from related species; PCR primers (see
CC   AAT90764-65) based on internal peptides P2 and P3 were used to
CC   isolate mouse and human neurturin cDNA clones.  Neurturin is a
CC   member of the glial-derived neurotrophic factor (GDNF) family.
CC   A novel member of the GDNF-neurturin growth factor family,
CC   persephin (see AAW30064-68), was subsequently identified that can
CC   be used in claimed methods for preventing or treating neuronal
CC   degeneration, haematopoietic cell degeneration and cardiac muscle
CC   degeneration or insufficiency.
XX
SQ   Sequence   10 AA;

  Query Match             40.0%;  Score 24;  DB 18;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 5e+02;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 HTITKLNA 11
          ||: :|:|
Db      2 htlqelsa 9
```

```
RESULT 15
AAW13715
ID   AAW13715 standard; Peptide; 10 AA.
XX
AC   AAW13715;
XX
DT   09-FEB-1998  (first entry)
XX
DE   Internal neurturin fragment P3.
XX
KW   Neurturin; human; haematopoietic cell; neuronal cell; stem cell; NT gene;
KW   neurodegenerative disease; peripheral neuropathy; nervous system tumour;
KW   amyotrophic lateral sclerosis; Alzheimer's disease; Parkinson's disease;
KW   Huntingdon's disease; ischaemic stroke; acute brain injury; basopaenia;
KW   acute spinal cord injury; multiple sclerosis; eosinopaenia; lymphopaenia;
KW   monocytopaenia; neutropaenia; anaemia; thromobocytopaenia; neuroblastoma;
KW   antibody; obesity; therapy; hamster.
XX
OS   Cricetulus griseus.
XX
PN   WO9708196-A1.
XX
PD   06-MAR-1997.
XX
PF   27-AUG-1996;   96WO-US14065.
XX
PR   28-AUG-1995;   95US-0519777.
XX
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Johnson EM,  Kotzbauer PT,  Lampe PA,  Milbrandt JD;
XX
DR   WPI; 1997-179176/16.
XX
PT   A novel growth factor Neurturin  - used to treat neuro-degenerative
PT   and haematopoietic cell degeneration diseases, e.g. Alzheimer's
PT   disease and eosinopenia
XX
PS   Example 5; Page 8; 206pp; English.
XX
CC   AAW13712-W13715 represent internal fragments of neurturin (NT). These
CC   sequences, and the N-terminal fragment shown in AAW13712, were isolated
CC   from chinese hamster ovary (CHO) cells. These sequences were used to
CC   produce degenerate primers (see AAT61481 and AAT61482) which were used
CC   to isolate the human and mouse NT coding sequences (see AAT61468 and
CC   AAT61469 respectively). NT promotes the growth and differentiation of
CC   haematopoietic and neuronal cells, and their stem cells. The NT gene and
CC   protein are used to prevent or treat neurodegenerative diseases e.g.
CC   peripheral neuropathy, amyotrophic lateral sclerosis, Alzheimer's
CC   disease, Parkinson's disease, Huntingdon's disease, ischaemic stroke,
CC   acute brain injury, acute spinal cord injury, nervous system tumours,
CC   multiple sclerosis and infection; and haematopoietic cell degenerative
CC   diseases, e.g. eosinopaenia, basopaenia, lymphopaenia, monocytopaenia,
CC   neutropaenia, anaemia, thromobocytopaenia and stem cell insufficiencies.
CC   The NT protein and gene are also useful to treat neuroblastomas.
CC   Antibodies against NT and oligonucleotides (used as either probes or
CC   primers, corresponding to an exon of pre-pro-NT gene or flanking a
CC   target sequence) can be used for detecting NT in a sample or detecting
CC   mutations in the NT gene. Antisense sequences of the NT gene are used to
CC   treat diseases promoted by NT expression e.g. obesity.
XX
SQ   Sequence   10 AA;

  Query Match             40.0%;  Score 24;  DB 18;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 5e+02;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 HTITKLNA 11
           ||: :|:|
Db       2 htlqelsa 9
```

```
RESULT 16
AAY16636
ID   AAY16636 standard; Peptide; 10 AA.
XX
AC   AAY16636;
XX
DT   17-AUG-1999  (first entry)
XX
DE   WO9914235 Seq ID No: 6.
XX
KW   Growth factor; GF; persephin; neuron growth; cellular degeneration;
KW   peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke;
KW   Alzheimer's disease; Parkinson's disease; Huntington's disease; trauma;
KW   brain injury; spinal cord injury; nervous system tumour; infection;
KW   multiple sclerosis; cardiac muscle degeneration; injury;neurotoxin;
KW   metabolic disease; diabetes; renal dysfunction; neurturin.
XX
OS   Mus sp.
XX
PN   WO9914235-A1.
XX
PD   25-MAR-1999.
XX
PF   15-SEP-1998;   98WO-US19163.
XX
PR   16-SEP-1997;   97US-0931858.
XX
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Desauvage F,  Johnson EM,  Klein R,  Kotzbauer PT;
PI   Lampe PA,  Milbrandt JD;
XX
DR   WPI; 1999-244023/20.
XX
PT   New isolated persephin growth factor nucleic acids used to, e.g.
PT   promote neuronal growth
XX
PS   Example 5; Page 125; 222pp; English.
XX
CC   The invention relates to a novel isolated and purified growth factor (GF)
CC   that comprises persephin or a fragment or a conservatively substituted
CC   variant. The persephin GF polypeptides can promote the survival and
CC   growth of neurons and non-neuronal cells. The persephin GF polypeptides
CC   or polynucleotides can be used for preventing or treating cellular
CC   degeneration or insufficiency, e.g. neuronal degeneration resulting from
CC   peripheral neuropathy, amyotrophic lateral sclerosis, Alzheimer's
CC   disease, Parkinson's disease, Huntington's disease, ischemic stroke,
CC   acute brain injury, acute spinal cord injury, nervous system tumours,
CC   multiple sclerosis, or infection, hematopoietic cell degeneration or
CC   insufficiency resulting from eosinopenia, anemias, thrombocytopenia, or
CC   stem-cell insufficiences, cardiac muscle degeneration or insufficiency
CC   resulting from cardiomyopathy or congestive heart failure. They can also
CC   be used for treating e.g. peripheral nerve trauma or injury, exposure to
CC   neurotoxins, metabolic diseases such as diabetes or renal dysfunctions
CC   and damage caused by infectious agents. The GF can also be used for
CC   promoting the growth and/or differentiation of a cell in a culture
CC   medium. The antisense polynucleotides can be used for treating a disease
CC   condition mediated by expression of persephin by a population of cells.
CC   The products can also be used for detection and diagnosis.
XX
SQ   Sequence   10 AA;

  Query Match             40.0%;  Score 24;  DB 20;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 5e+02;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 HTITKLNA 11
           ||: :|:|
Db       2 htlqelsa 9
```

```
RESULT 17
AAM42846
ID   AAM42846 standard; Peptide; 10 AA.
XX
AC   AAM42846;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Mycoplasma genitalium intermolecular complementary peptide, SEQ ID 155.
XX
KW   Mycoplasma genitalium; complementary peptide; ligand;
KW   protein-protein interaction; drug design; intermolecular;
KW   intramolecular.
XX
XX
OS   Mycoplasma genitalium.
XX
PN   WO200142278-A2.
XX
PD   14-JUN-2001.
XX
PF   13-DEC-2000; 2000WO-GB04778.
XX
PR   13-DEC-1999;   99GB-0029466.
XX
PA   (PROT-) PROTEOM LTD.
XX
PI   Roberts GW,  Heal JR;
XX
DR   WPI; 2001-514238/56.
XX
PT   Complementary peptide ligands as reagents and drugs for drug discovery
PT   programs and as lead ligands to facilitate drug design and development,
PT   are generated from microbial genome sequences  -
XX
PS   Example 2; Page 71; 161pp; English.
XX
CC   The present sequence is one of a large number of complementary peptide
CC   ligands generated from Mycoplasma genitalium genome sequences. These
CC   specific complementary peptides interact with their relevant target
CC   proteins encoded by the microbial genome. They are capable of
CC   antagonising or agonising specific interaction of a protein with
CC   another protein or receptor and are thus useful as reagents and drugs,
CC   and as lead ligands to facilitate drug design and development. They
CC   are useful as tools for functional genomic studies, reagents for the
CC   configuration of high-throughput screens, as a starting point for
CC   medicinal chemistry manipulation, for peptide mimetics and as
CC   therapeutic agents. The analysis and acquisition of peptide sequences
CC   facilitates understanding of protein-protein interactions. The method
CC   allows for analysis of an entire database at a time, thus overcoming
CC   sampling problems. The set of complementary peptides includes both
CC   intermolecular (between proteins) and intermolecular (within a
CC   protein) sequences.
XX
SQ   Sequence   10 AA;

  Query Match             40.0%;  Score 24;  DB 22;  Length 10;
  Best Local Similarity   83.3%;  Pred. No. 5e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 TITKLN 10
           || |||
Db       2 tinkln 7


RESULT 18
AAR87683
ID   AAR87683 standard; peptide; 11 AA.
XX
AC   AAR87683;
XX
DT   10-MAY-1996  (first entry)
XX
DE   h-TNF (70-80) analogue.
XX
KW   TNF; tumour necrosis factor; neutrophil; superoxide; inflammation;
KW   AIDS, cancer; rheumatoid arthritis; ARDS; cystic fibrosis;
KW   tuberculosis; TB; Crohn's disease; sarcoidosis; leprosy; syphilis;
KW   malaria; adult respiratory distress syndrome.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "this residue is optionally N-acylated or
FT                          glycosylated"
FT   Modified-site   11
FT                   /note= "this residue is optionally in amide or
FT                          ester form and may be glycosylated"
XX
PN   AU9344664-A.
XX
PD   02-MAR-1995.
XX
PF   12-MAR-1991;   91AU-0074762.
XX
PR   12-MAR-1991;   91AU-0074762.
PR   13-AUG-1993;   93AU-0044664.
XX
PA   (PEPT-) PEPTIDE TECHNOLOGY LTD.
XX
PI   Ferrante A,  Rathjen DA,  Widmer F;
XX
DR   WPI; 1995-115742/16.
XX
PT   Novel peptides which stimulate neutrophils - are fragments of
PT   tumour necrosis factor
XX
PS   Claim 4; Page 32; 51pp; English.
XX
CC   New peptides are provided which prime neutrophils for superoxide
CC   production and an enhanced respiratory burst following treatment
CC   with N-formyl-L-methionine-L-leucyl-L-phenylalanine. The peptides are
CC   described generically. They are fragments from the region of amino acids
CC   54 to 94 of human TNF, or analogues of such fragments. Their
CC   neutrophil-stimulating activity can be used in the treatment of diseases
CC   such as AIDS, cancer, rheumatoid arthritis, ARDS, cystic fibrosis,
CC   tuberculosis, Crohn's disease, sarcoidosis, leprosy, syphilis and
CC   malaria. The present sequence is one of 6 specific peptides having the
CC   above activity and which fit the the generic formula.
XX
SQ   Sequence   11 AA;

  Query Match             40.0%;  Score 24;  DB 16;  Length 11;
  Best Local Similarity   80.0%;  Pred. No. 5.6e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 LTHTI 6
           :||||
Db       7 ithti 11


RESULT 19
AAW35449
ID   AAW35449 standard; peptide; 11 AA.
XX
AC   AAW35449;
XX
DT   22-APR-1998  (first entry)
XX
DE   Cytokine derived peptide from TNF (70-80).
XX
KW   T-cell stimulatory peptide; immunogen; non-dendritic; carrier; tumour;
KW   scaffold; inhibition; metastasis; wound healing; solid phase.
```

XX
OS Unidentified.
XX
PN WO9738011-A1.
XX
PD 16-OCT-1997.
XX
PF 03-APR-1997; 97WO-DE00146.
XX
PR 03-APR-1996; 96DK-0000398.
XX
PA (PEPR-) PEPRESEARCH AS.
XX
PI Heegaard PMH, Jakobsen PH;
XX
DR WPI; 1997-512645/47.
XX
PT Non-dendritic peptide carrier linked to a solid phase - useful as a
PT diagnostic agent and as a scaffold for production of chemical
PT derivatives
XX
PS Claim 30; Page 199; 262pp; English.
XX
CC A non-dendritic peptide carrier (A) has been developed which is coupled
CC through a linker to a solid phase, forming a complex of (A)-solid phase.
CC Where (A) comprises 10-50 amino acids capable of forming a secondary
CC structure in a benign buffer after liberation from the solid phase, and
CC further the (A)-solid phase complex comprises an immunogenic substance
CC and/or an immune mediator coupled on (A). The present sequence
CC represents a specifically claimed cytokine derived peptide from the
CC present invention. An (A)-solid phase complex can be used as a scaffold
CC for the production of chemical derivatives, characterised by covalently
CC attaching molecules at attachment points. Alternatively (A) is used as
CC a scaffold-peptide for the incorporation into an Immunostimulating
CC Complex (Iscom) resulting an (A)-Iscom complex which is used for the
CC chemical coupling of antigenic substances in an aqueous solution by
CC conjugation. (A) derivatised with one or more peptides having
CC fibronectin-, laminin- or vitronectin-like binding activities can be
CC used for the promotion of cell-attachment to plastic surfaces, in
CC particular to inhibit tumour growth and metastasis, and for promotion
CC of wound healing. Also a derivatised (A) can be used for the selection
CC of specifically-binding aptamers or as a diagnostic agent. Such
CC diagnostic-(A) molecules could be used to detect molecules derived from
CC or indicative of pregnancy or of a disease, such as an infectious,
CC autoimmune or cancerous disease.
XX
SQ Sequence 11 AA;

Query Match 40.0%; Score 24; DB 18; Length 11;
Best Local Similarity 80.0%; Pred. No. 5.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LTHTI 6
:||||
Db 7 ithti 11

RESULT 20
AAW47911
ID AAW47911 standard; peptide; 11 AA.
XX
AC AAW47911;
XX
DT 09-JUN-1998 (first entry)
XX
DE Human tumour necrosis factor derived peptide 1168.
XX
KW Human; tumour necrosis factor; neutrophil; monocyte; macrophage;
KW stimulatory activity; immunosuppressant; cancer; infection.
XX
OS Synthetic.
OS Homo sapiens.

XX
PN WO9748725-A1.
XX
PD 24-DEC-1997.
XX
PF 20-JUN-1997; 97WO-AU00395.
XX
PR 29-OCT-1996; 96AU-0003309.
PR 21-JUN-1996; 96AU-0000610.
PR 06-SEP-1996; 96AU-0002165.
XX
PA (PEPT-) PEPTECH LTD.
XX
PI Mack PO, Rathjen DA, Sleigh JM, Widmer F;
XX
DR WPI; 1998-063077/06.
XX
PT Tumour necrosis factor derived peptide(s) having neutrophil and-or
PT monocyte-macrophage stimulatory activity - used for treating e.g.
PT infections, immunosuppression or cancers
XX
PS Claim 4; Page 44; 65pp; English.
XX
CC The present sequence represents a peptide derived from a human tumour
CC necrosis factor peptide. The peptide can have neutrophil and/or
CC monocyte/macrophage stimulatory activity. The peptide is derived from
CC the sequence of tumour necrosis factor (TNF)-derived peptide 419
CC (PSTHVLITHTI; see AU74762/91 and 44664/93). It can have improved
CC properties such as increased potency, extended in vivo half life or,
CC particularly, specificity of action. It can have neutrophil stimulatory
CC activity (class 1), equal neutrophil and monocyte/macrophage stimulatory
CC activity (class 2), or preferentially enhanced monocyte/macrophage
CC stimulatory activity (class 3). It can be used for the treatment or
CC prevention of infections. In particular, it can be used for treating
CC AIDS, cancer, diabetes, nosocomial infection, tuberculosis, cystic
CC fibrosis, community acquired pneumonia, meningitis, Mycobacteria,
CC Chlamydia, Brucellae, Francisella, Pasteurellosis, Legionellosis,
CC Histoplasmosis, Listeriosis, Pneumocystis carnii, Trypanosoma cruzi,
CC coccidian parasitical infection, an inherited primary neutropenic
CC disorder, an inherited primary defect of phagocytic cell function, an
CC inherited secondary defect of phagocytic cell function, an acquired
CC defect of phagocytic cell function, immunosuppression due to the
CC administration of immunosuppressive drugs, and other bacterial, fungal,
CC viral or protozoan infection, infectious mononucleosis, paroxysomal
CC nocturnal, haemoglobinuria, leukaemia, lymphoma, myelofibrosis, or graft
CC versus host disease.
XX
SQ Sequence 11 AA;

Query Match 40.0%; Score 24; DB 19; Length 11;
Best Local Similarity 80.0%; Pred. No. 5.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LTHTI 6
:||||
Db 7 ithti 11

Search completed: July 8, 2002, 11:42:57
Job time: 229 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 8, 2002, 11:40:58 ; Search time 21.49 Seconds
(without alignments)
13.639 Million cell updates/sec

Title: US-09-461-061A-2
Perfect score: 60
Sequence: 1 TLTHTITKLNAE 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 83765

Minimum DB seq length: 0
Maximum DB seq length: 12

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	31	51.7	9	1	US-08-107-235-13	Sequence 13, Appl
2	31	51.7	9	3'	US-08-714-960B-13	Sequence 13, Appl
3	26	43.3	9	1	US-08-259-550A-21	Sequence 21, Appl
4	26	43.3	10	4	US-09-177-249-131	Sequence 131, App
5	26	43.3	11	1	US-08-107-235-15	Sequence 15, Appl
6	26	43.3	11	1	US-08-107-235-17	Sequence 17, Appl
7	26	43.3	11	3	US-08-916-443A-1	Sequence 1, Appli
8	26	43.3	11	3	US-08-714-960B-15	Sequence 15, Appl
9	26	43.3	11	3	US-08-714-960B-17	Sequence 17, Appl
10	24	40.0	10	1	US-08-519-777-6	Sequence 6, Appli
11	24	40.0	10	1	US-08-742-035-6	Sequence 6, Appli
12	24	40.0	10	2	US-08-777-019-6	Sequence 6, Appli
13	24	40.0	10	2	US-08-777-143-6	Sequence 6, Appli
14	24	40.0	10	3	US-08-775-414-6	Sequence 6, Appli
15	24	40.0	10	4	US-08-931-858E-6	Sequence 6, Appli
16	24	40.0	10	4	US-08-981-739-6	Sequence 6, Appli
17	24	40.0	11	1	US-08-107-235-18	Sequence 18, Appl
18	24	40.0	11	3	US-08-714-960B-18	Sequence 18, Appl
19	23	38.3	10	2	US-08-025-321C-1	Sequence 1, Appli
20	23	38.3	12	4	US-09-461-697-138	Sequence 138, App
21	22	36.7	5	4	US-08-751-344B-29	Sequence 29, Appl
22	22	36.7	7	4	US-08-640-737-38	Sequence 38, Appl
23	22	36.7	11	4	US-09-306-756-3	Sequence 3, Appli
24	22	36.7	12	3	US-09-188-579-33	Sequence 33, Appl
25	22	36.7	12	4	US-09-315-444-33	Sequence 33, Appl
26	22	36.7	12	4	US-09-268-347-17	Sequence 17, Appl
27	22	36.7	12	4	US-09-382-689A-6	Sequence 6, Appli
28	22	36.7	12	4	US-09-261-182-9	Sequence 9, Appli
29	21	35.0	9	4	US-09-258-754-146	Sequence 146, App
30	21	35.0	9	4	US-09-042-107-146	Sequence 146, App
31	21	35.0	11	1	US-07-657-769B-43	Sequence 43, Appl
32	21	35.0	11	1	US-07-789-184-104	Sequence 104, App
33	21	35.0	11	1	US-07-789-184-215	Sequence 215, App
34	21	35.0	11	1	US-07-789-184-218	Sequence 218, App
35	21	35.0	11	1	US-08-475-263-104	Sequence 104, App
36	21	35.0	11	1	US-08-475-263-215	Sequence 215, App
37	21	35.0	11	1	US-08-475-263-218	Sequence 218, App
38	21	35.0	11	1	US-08-485-886-104	Sequence 104, App
39	21	35.0	11	1	US-08-485-886-215	Sequence 215, App
40	21	35.0	11	1	US-08-485-886-218	Sequence 218, App
41	21	35.0	11	2	US-08-477-362-104	Sequence 104, App
42	21	35.0	11	2	US-08-477-362-215	Sequence 215, App
43	21	35.0	11	2	US-08-477-362-218	Sequence 218, App
44	21	35.0	11	2	US-08-477-134-104	Sequence 104, App
45	21	35.0	11	2	US-08-477-134-215	Sequence 215, App

ALIGNMENTS

RESULT 1
US-08-107-235-13
; Sequence 13, Application US/08107235
; Patent No. 5587457
; GENERAL INFORMATION:
; APPLICANT: Rathjen, Deborah A
; APPLICANT: Ferrante, Antonio
; APPLICANT: Widmer, Fred
; TITLE OF INVENTION: Neutrophil Stimulating Peptides
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 S. Wacker Dr.
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/107,235
; FILING DATE: 16-AUG-1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/930,415
; FILING DATE: 12-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,622A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..9
; OTHER INFORMATION: /note= "PEPTIDE 393 (76-84)"
US-08-107-235-13

Query Match 51.7%; Score 31; DB 1; Length 9;

```
  Best Local Similarity   62.5%;  Pred. No. 1.7e+05;
  Matches    5;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 LTHTITKL 9
           |||||:::
Db       1 LTHTISRI 8


RESULT   2
US-08-714-960B-13
; Sequence 13, Application US/08714960B
; Patent No. 6121237
;  GENERAL INFORMATION:
;    APPLICANT:  RATHJEN, Deborah A
;    APPLICANT:  FERRANTE, Antonio
;    TITLE OF INVENTION:  Neutrophil Stimulating Peptides
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BANNER & WITCOFF, LTD.
;      STREET:  10 S. Wacker Drive, Suite 3000
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  1.44 Mb storage diskette, 3.50 inch
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  IBM compatible PC/MS-DOS
;      SOFTWARE:  WordPerfect version 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/714,960B
;      FILING DATE:  17-SEP-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  AU PJ9065
;      FILING DATE:  12-MAR-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/AU91/00086
;      FILING DATE:  12-MAR-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/930,415
;      FILING DATE:  09-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/107,235
;      FILING DATE:  16-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Resis, Robert H.
;      REGISTRATION NUMBER:  32,168
;      REFERENCE/DOCKET NUMBER:  92,622-B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 715-1000
;      TELEFAX:  (312) 715-1234
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Peptide
;      LOCATION:  1..9
;      OTHER INFORMATION:  /note= "PEPTIDE 393 (76-84)"
US-08-714-960B-13


  Query Match           51.7%;  Score 31;  DB 3;  Length 9;
  Best Local Similarity   62.5%;  Pred. No. 1.7e+05;
  Matches    5;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 LTHTITKL 9
           |||||:::
Db       1 LTHTISRI 8


RESULT   3
US-08-259-550A-21
; Sequence 21, Application US/08259550A
; Patent No. 5776892
;  GENERAL INFORMATION:
;    APPLICANT:  Counts, David F.
;    APPLICANT:  Duff, Ronald G.
;    TITLE OF INVENTION:  Anti-Inflammatory Peptides
;    NUMBER OF SEQUENCES:  91
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PENNIE & EDMONDS
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/259,550A
;      FILING DATE:  16-JUN-1994
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Misrock, S. Leslie
;      REGISTRATION NUMBER:  18,872
;      REFERENCE/DOCKET NUMBER:  7142-011
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-8864/9741
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-259-550A-21


  Query Match           43.3%;  Score 26;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 LTHTI 6
           |||||
Db       5 LTHTI 9


RESULT   4
US-09-177-249-131
; Sequence 131, Application US/09177249
; Patent No. 6229064
; GENERAL INFORMATION:
;  APPLICANT: Fischer, Robert L.
;  APPLICANT: Ohad, Nir
;  APPLICANT: Kiyosue, Tomohiro
;  APPLICANT: Yadegari, Ramin
;  APPLICANT: Margossian, Linda
;  APPLICANT: Harada, John
;  APPLICANT: Goldberg, Robert B.
;  APPLICANT: The Regents of the University of California
;  TITLE OF INVENTION: Nucleic Acids That Control Seed and Fruit
;  TITLE OF INVENTION: Development in Plants
;  FILE REFERENCE: 02307O-086120US
;  CURRENT APPLICATION NUMBER: US/09/177,249
;  CURRENT FILING DATE: 1998-10-22
```

```
;  EARLIER APPLICATION NUMBER: US 09/071,838
;  EARLIER FILING DATE: 1998-05-01
;  NUMBER OF SEQ ID NOS: 324
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 131
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Arabidopsis sp.
US-09-177-249-131


  Query Match           43.3%;  Score 26;  DB 4;  Length 10;
  Best Local Similarity  55.6%;  Pred. No. 63;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      4 HTITKLNAE 12
          ||: | | |
Db      1 HTLFKKNLE 9


RESULT   5
US-08-107-235-15
; Sequence 15, Application US/08107235
; Patent No. 5587457
;  GENERAL INFORMATION:
;    APPLICANT:  Rathjen, Deborah A
;    APPLICANT:  Ferrante, Antonio
;    APPLICANT:  Widmer, Fred
;    TITLE OF INVENTION:  Neutrophil Stimulating Peptides
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Allegretti & Witcoff, Ltd.
;      STREET:  10 S. Wacker Dr.
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/107,235
;      FILING DATE:  16-AUG-1993
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/930,415
;      FILING DATE:  12-MAR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McDonnell, John J
;      REGISTRATION NUMBER:  26,949
;      REFERENCE/DOCKET NUMBER:  92,622A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-715-1000
;      TELEFAX:  312-715-1234
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Peptide
;      LOCATION:  1..11
;      OTHER INFORMATION:  /note= "PEPTIDE 395 (70-80)"
US-08-107-235-15


  Query Match           43.3%;  Score 26;  DB 1;  Length 11;
  Best Local Similarity  100.0%;  Pred. No. 70;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 LTHTI 6
          |||||
Db      7 LTHTI 11


RESULT   6
US-08-107-235-17
; Sequence 17, Application US/08107235
; Patent No. 5587457
;  GENERAL INFORMATION:
;    APPLICANT:  Rathjen, Deborah A
;    APPLICANT:  Ferrante, Antonio
;    APPLICANT:  Widmer, Fred
;    TITLE OF INVENTION:  Neutrophil Stimulating Peptides
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Allegretti & Witcoff, Ltd.
;      STREET:  10 S. Wacker Dr.
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/107,235
;      FILING DATE:  16-AUG-1993
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/930,415
;      FILING DATE:  12-MAR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McDonnell, John J
;      REGISTRATION NUMBER:  26,949
;      REFERENCE/DOCKET NUMBER:  92,622A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-715-1000
;      TELEFAX:  312-715-1234
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Peptide
;      LOCATION:  1..11
;      OTHER INFORMATION:  /note= "PEPTIDE 418"
US-08-107-235-17


  Query Match           43.3%;  Score 26;  DB 1;  Length 11;
  Best Local Similarity  100.0%;  Pred. No. 70;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 LTHTI 6
          |||||
Db      7 LTHTI 11


RESULT   7
US-08-916-443A-1
; Sequence 1, Application US/08916443A
; Patent No. 6001986
;  GENERAL INFORMATION:
;    APPLICANT:  Yong Sig KIM
;    APPLICANT:  Sun Chung PARK
;    APPLICANT:  Soo Kyung OH
;    APPLICANT:  Hosull LEE
;    APPLICANT:  Jeong Woo CHO
```

```
;    APPLICANT:  Chang H. CHUNG
;    TITLE OF INVENTION:  Antiviral Proteins, Amarandin 1 and 2, from
;    TITLE OF INVENTION:  Amaranthus Viridis, DNAs Encoding Therefrom
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kenyon & Kenyon
;      STREET:  1025 Connecticut Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3+ Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS 6.2
;      SOFTWARE:  WordPerfect 6.1 Windows
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/916,443A
;      FILING DATE:  22 AUG 1997
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Toffenetti, Judith L.
;      REGISTRATION NUMBER:  39,048
;      REFERENCE/DOCKET NUMBER:  1942/18
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-429-1776
;      TELEFAX:  202-429-0796
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-916-443A-1


  Query Match             43.3%;  Score 26;  DB 3;  Length 11;
  Best Local Similarity   71.4%;  Pred. No. 70;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 LTHTITK 8
           || |:||
Db       3 LTFTVTK 9


RESULT   8
US-08-714-960B-15
; Sequence 15, Application US/08714960B
; Patent No. 6121237
;  GENERAL INFORMATION:
;    APPLICANT:  RATHJEN, Deborah A
;    APPLICANT:  FERRANTE, Antonio
;    TITLE OF INVENTION:  Neutrophil Stimulating Peptides
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BANNER & WITCOFF, LTD.
;      STREET:  10 S. Wacker Drive, Suite 3000
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  1.44 Mb storage diskette, 3.50 inch
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  IBM compatible PC/MS-DOS
;      SOFTWARE:  WordPerfect version 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/714,960B
;      FILING DATE:  17-SEP-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  AU PJ9065
;      FILING DATE:  12-MAR-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/AU91/00086
;      FILING DATE:  12-MAR-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/930,415
;      FILING DATE:  09-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/107,235
;      FILING DATE:  16-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Resis, Robert H.
;      REGISTRATION NUMBER:  32,168
;      REFERENCE/DOCKET NUMBER:  92,622-B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 715-1000
;      TELEFAX:  (312) 715-1234
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Peptide
;      LOCATION:  1..11
;      OTHER INFORMATION:  /note= "PEPTIDE 395 (70-80)"
US-08-714-960B-15


  Query Match             43.3%;  Score 26;  DB 3;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 70;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 LTHTI 6
             |||||
Db       7 LTHTI 11


RESULT   9
US-08-714-960B-17
; Sequence 17, Application US/08714960B
; Patent No. 6121237
;  GENERAL INFORMATION:
;    APPLICANT:  RATHJEN, Deborah A
;    APPLICANT:  FERRANTE, Antonio
;    TITLE OF INVENTION:  Neutrophil Stimulating Peptides
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BANNER & WITCOFF, LTD.
;      STREET:  10 S. Wacker Drive, Suite 3000
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  1.44 Mb storage diskette, 3.50 inch
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  IBM compatible PC/MS-DOS
;      SOFTWARE:  WordPerfect version 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/714,960B
;      FILING DATE:  17-SEP-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  AU PJ9065
;      FILING DATE:  12-MAR-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/AU91/00086
;      FILING DATE:  12-MAR-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/930,415
```

```
;      FILING DATE:  09-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/107,235
;      FILING DATE:  16-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Resis, Robert H.
;      REGISTRATION NUMBER:  32,168
;      REFERENCE/DOCKET NUMBER:  92,622-B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 715-1000
;      TELEFAX:  (312) 715-1234
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Peptide
;      LOCATION:  1..11
;      OTHER INFORMATION:  /note= "PEPTIDE 418"
US-08-714-960B-17


  Query Match             43.3%;  Score 26;  DB 3;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 70;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 LTHTI 6
           |||||
Db       7 LTHTI 11


RESULT  10
US-08-519-777-6
; Sequence 6, Application US/08519777
; Patent No. 5739307
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON JR., EUGENE M.
;    APPLICANT:  MILBRANDT, JEFFREY D.
;    APPLICANT:  KOTZBAUER, PAUL T.
;    APPLICANT:  LAMPE, PATRICIA A.
;    TITLE OF INVENTION:  NEURTURIN AND RELATED GROWTH FACTORS
;    NUMBER OF SEQUENCES:  78
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROGERS, HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  US
;      ZIP:  63105-1817
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/519,777
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  953095
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-519-777-6


  Query Match             40.0%;  Score 24;  DB 1;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 HTITKLNA 11
           ||: :|:|
Db       2 HTLQELSA 9


RESULT  11
US-08-742-035-6
; Sequence 6, Application US/08742035
; Patent No. 5747655
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON JR., EUGENE M.
;    APPLICANT:  MILBRANDT, JEFFREY D.
;    APPLICANT:  KOTZBAUER, PAUL T.
;    APPLICANT:  LAMPE, PATRICIA A.
;    TITLE OF INVENTION:  NEURTURIN AND RELATED GROWTH FACTORS
;    NUMBER OF SEQUENCES:  78
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROGERS, HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  US
;      ZIP:  63105-1817
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/742,035
;      FILING DATE:  01-NOV-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/519,777
;      FILING DATE:  28-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  953095
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-742-035-6


  Query Match             40.0%;  Score 24;  DB 1;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 HTITKLNA 11
           ||: :|:|
Db       2 HTLQELSA 9


RESULT  12
US-08-777-019-6
; Sequence 6, Application US/08777019
; Patent No. 5817622
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  JOHNSON JR., EUGENE M.
;    APPLICANT:  MILBRANDT, JEFFREY D.
;    APPLICANT:  KOTZBAUER, PAUL T.
;    APPLICANT:  LAMPE, PATRICIA A.
;    TITLE OF INVENTION:  NEURTURIN AND RELATED GROWTH FACTORS
;    NUMBER OF SEQUENCES:  78
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROGERS, HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  US
;      ZIP:  63105-1817
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/777,019
;      FILING DATE:  30-DEC-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/519,777
;      FILING DATE:  28-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  953095
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-777-019-6


  Query Match             40.0%;  Score 24;  DB 2;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 HTITKLNA 11
           ||: :|:|
Db       2 HTLQELSA 9


RESULT  13
US-08-777-143-6
; Sequence 6, Application US/08777143
; Patent No. 5843914
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON JR., EUGENE M.
;    APPLICANT:  MILBRANDT, JEFFREY D.
;    APPLICANT:  KOTZBAUER, PAUL T.
;    APPLICANT:  LAMPE, PATRICIA A.
;    TITLE OF INVENTION:  NEURTURIN AND RELATED GROWTH FACTORS
;    NUMBER OF SEQUENCES:  78
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROGERS, HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  US
;      ZIP:  63105-1817
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/777,143
;      FILING DATE:  30-DEC-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/519,777
;      FILING DATE:  28-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  953095
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-777-143-6


  Query Match             40.0%;  Score 24;  DB 2;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 HTITKLNA 11
           ||: :|:|
Db       2 HTLQELSA 9


RESULT  14
US-08-775-414-6
; Sequence 6, Application US/08775414
; Patent No. 6090778
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON JR., EUGENE M.
;    APPLICANT:  MILBRANDT, JEFFREY D.
;    APPLICANT:  KOTZBAUER, PAUL T.
;    APPLICANT:  LAMPE, PATRICIA A.
;    TITLE OF INVENTION:  NEURTURIN AND RELATED GROWTH FACTORS
;    NUMBER OF SEQUENCES:  90
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  US
;      ZIP:  63105-1817
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/775,414
;      FILING DATE:  31-DEC-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965805
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-775-414-6
```

```
  Query Match             40.0%;  Score 24;  DB 3;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 HTITKLNA 11
          ||: :|:|
Db      2 HTLQELSA 9


RESULT  15
US-08-931-858E-6
; Sequence 6, Application US/08931858E
; Patent No. 6222022
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON, EUGENE M
;    APPLICANT:  MILBRANDT, JEFFREY D
;    APPLICANT:  KOTZBAUER, PAUL T
;    APPLICANT:  LAMPE, PATRICIA A
;    APPLICANT:  KLEIN, ROBERT
;    APPLICANT:  DESAUVAGE, FRED
;    TITLE OF INVENTION:  PERSEPHIN AND RELATED GROWTH FACTOR
;    NUMBER OF SEQUENCES:  239
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BOULEVARD, SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MO
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/931,858E
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  971486
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  314-727-5188
;      TELEFAX:  314-727-6092
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-931-858E-6


  Query Match             40.0%;  Score 24;  DB 4;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 HTITKLNA 11
          ||: :|:|
Db      2 HTLQELSA 9


RESULT  16
US-08-981-739-6
; Sequence 6, Application US/08981739
; Patent No. 6232449
;   GENERAL INFORMATION:
;        APPLICANT: JOHNSON JR., EUGENE M.
;                   MILBRANDT, JEFFREY D.
;                   KOTZBAUER, PAUL T.
;                   LAMPE, PATRICIA A.
;        TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTORS
;        NUMBER OF SEQUENCES: 176
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: HOWELL & HAFERKAMP, L.C.
;             STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
;             CITY: ST. LOUIS
;             STATE: MISSOURI
;             COUNTRY: US
;             ZIP: 63105-1817
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/981,739
;             FILING DATE: 31-Aug-1998
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: PCT/US97/03461
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: HOLLAND, DONALD R.
;             REGISTRATION NUMBER: 35,197
;             REFERENCE/DOCKET NUMBER: 976163
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (314) 727-5188
;             TELEFAX: (314) 727-6092
;   INFORMATION FOR SEQ ID NO: 6:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 10 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: <Unknown>
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-08-981-739-6


  Query Match             40.0%;  Score 24;  DB 4;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 HTITKLNA 11
          ||: :|:|
Db      2 HTLQELSA 9


RESULT  17
US-08-107-235-18
; Sequence 18, Application US/08107235
; Patent No. 5587457
;  GENERAL INFORMATION:
;    APPLICANT:  Rathjen, Deborah A
;    APPLICANT:  Ferrante, Antonio
;    APPLICANT:  Widmer, Fred
;    TITLE OF INVENTION:  Neutrophil Stimulating Peptides
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Allegretti & Witcoff, Ltd.
;      STREET:  10 S. Wacker Dr.
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US/08/107,235
;      FILING DATE:  16-AUG-1993
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/930,415
;      FILING DATE:  12-MAR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McDonnell, John J
;      REGISTRATION NUMBER:  26,949
;      REFERENCE/DOCKET NUMBER:  92,622A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-715-1000
;      TELEFAX:  312-715-1234
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Peptide
;      LOCATION:  1..11
;      OTHER INFORMATION:  /note= "PEPTIDE 419"
US-08-107-235-18


  Query Match             40.0%;  Score 24;  DB 1;  Length 11;
  Best Local Similarity   80.0%;  Pred. No. 1.6e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 LTHTI 6
           :||||
Db       7 ITHTI 11


RESULT  18
US-08-714-960B-18
; Sequence 18, Application US/08714960B
; Patent No. 6121237
;  GENERAL INFORMATION:
;    APPLICANT:  RATHJEN, Deborah A
;    APPLICANT:  FERRANTE, Antonio
;    TITLE OF INVENTION:  Neutrophil Stimulating Peptides
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BANNER & WITCOFF, LTD.
;      STREET:  10 S. Wacker Drive, Suite 3000
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  1.44 Mb storage diskette, 3.50 inch
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  IBM compatible PC/MS-DOS
;      SOFTWARE:  WordPerfect version 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/714,960B
;      FILING DATE:  17-SEP-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  AU PJ9065
;      FILING DATE:  12-MAR-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/AU91/00086
;      FILING DATE:  12-MAR-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/930,415
;      FILING DATE:  09-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/107,235
;      FILING DATE:  16-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Resis, Robert H.
;      REGISTRATION NUMBER:  32,168
;      REFERENCE/DOCKET NUMBER:  92,622-B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 715-1000
;      TELEFAX:  (312) 715-1234
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Peptide
;      LOCATION:  1..11
;      OTHER INFORMATION:  /note= "PEPTIDE 419"
US-08-714-960B-18


  Query Match             40.0%;  Score 24;  DB 3;  Length 11;
  Best Local Similarity   80.0%;  Pred. No. 1.6e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 LTHTI 6
           :||||
Db       7 ITHTI 11


RESULT  19
US-08-025-321C-1
; Sequence 1, Application US/08025321C
; Patent No. 5849560
;  GENERAL INFORMATION:
;    APPLICANT:  Abraham Ph.D., Carmela R.
;    TITLE OF INVENTION:  PROTEASES CAUSING ABNORMAL DEGREDATION
;    TITLE OF INVENTION:  OF AMYLOID BETA-PROTEIN PRECURSOR
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Choate, Hall & Stewart
;      STREET:  53 State Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/025,321C
;      FILING DATE:  26-FEB-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jarrell Ph.D., Brenda H.
;      REGISTRATION NUMBER:  39,223
;      REFERENCE/DOCKET NUMBER:  0079571-0034
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617 248 5000
;      TELEFAX:  617 248 4000
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  peptide
US-08-025-321C-1


  Query Match             38.3%;  Score 23;  DB 2;  Length 10;
  Best Local Similarity   44.4%;  Pred. No. 2.1e+02;
```

Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```
Qy      4 HTITKLNAE 12
          |:  |::||
Db      1 HSEVKMDAE 9
```

RESULT 20
US-09-461-697-138
; Sequence 138, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Puranam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; TITLE OF INVENTION: CELL DEATH
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/461,697
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 138
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-461-697-138

Query Match 38.3%; Score 23; DB 4; Length 12;
Best Local Similarity 83.3%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      7 TKLNAE 12
          |||| |
Db      7 TKLNNE 12
```

Search completed: July 8, 2002, 11:43:26
Job time: 148 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 8, 2002, 11:43:57 ; Search time 24.98 Seconds
(without alignments)
46.160 Million cell updates/sec

Title: US-09-461-061A-3
Perfect score: 56
Sequence: 1 IDNVKKARVQVV 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 1579

Minimum DB seq length: 0
Maximum DB seq length: 12

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	23	41.1	10	2	PS0220	ferredoxin--NADP+
2	22	39.3	10	2	S70251	nitrogenase (EC 1.
3	20	35.7	11	2	S33519	probable secreted
4	20	35.7	12	2	E64573	hypothetical prote
5	19	33.9	12	2	D28551	hypothetical prote
6	18	32.1	10	2	PA0116	ferredoxin--NADP+
7	17	30.4	7	2	A28709	phosphonoacetaldeh
8	17	30.4	11	2	JQ0395	hypothetical prote
9	16	28.6	10	2	S70722	65.4K GTP-binding
10	16	28.6	11	2	S19301	endo-1,4-beta-xyla
11	15	26.8	7	2	A15398	choline oxidase (E
12	15	26.8	10	2	PC2044	beta-Kirilowin - M
13	15	26.8	11	2	S21727	gamma-interferon-i
14	15	26.8	12	2	A60757	enterotoxin C-1 -
15	15	26.8	12	2	S18722	matK protein - bee
16	15	26.8	12	2	S65626	phosphorylase b -
17	14	25.0	6	2	B60110	repetitive protein
18	14	25.0	7	2	S19630	ribosomal protein
19	14	25.0	8	2	T10077	hypothetical prote
20	14	25.0	9	2	PS0253	glycine cleavage s
21	14	25.0	10	2	A32543	cardioexcitatory n
22	14	25.0	10	2	F44644	neurotoxin-associa
23	14	25.0	10	2	PQ0788	NADH dehydrogenase
24	14	25.0	10	2	PS0451	24K protein 4302 -
25	14	25.0	10	2	A42089	transcription fact
26	14	25.0	10	2	A56633	neomyosuppressin -
27	14	25.0	10	2	E86128	hypothetical prote
28	14	25.0	11	2	PH1343	Ig heavy chain DJ
29	14	25.0	11	2	PN0044	protein kinase C i
30	14	25.0	12	2	A39233	myosin heavy chain
31	14	25.0	12	2	S36902	Em protein - wheat
32	14	25.0	12	2	C61308	hemocyanin chain 6
33	14	25.0	12	2	A34858	proteinase E - bla
34	14	25.0	12	2	I77529	estrogen receptor
35	14	25.0	12	2	S21205	Ig heavy chain V r
36	14	25.0	12	2	I46922	gene Bota protein
37	13	23.2	6	2	PC4392	whey glycoprotein
38	13	23.2	7	2	A34818	vicilin 72K chain
39	13	23.2	8	2	A39892	P element, P cytot
40	13	23.2	8	2	S66296	Na+-transporting A
41	13	23.2	9	2	B39841	dextransucrase (EC
42	13	23.2	9	2	A39841	sucrose 3-glucosyl
43	13	23.2	9	2	S78426	52.5K protein - sp
44	13	23.2	10	2	PC2172	triacylglycerol li
45	13	23.2	10	2	C38925	seed storage prote

ALIGNMENTS

RESULT 1
PS0220
ferredoxin--NADP+ reductase (EC 1.18.1.2) - rice (fragment)
C;Species: Oryza sativa (rice)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-Feb-1997
C;Accession: PS0220
R;Uchiyama, Y.; Tsugita, A.
submitted to JIPID, August 1991
A;Reference number: PS0205
A;Accession: PS0220
A;Molecule type: protein
A;Residues: 1-10 <UCH>
C;Keywords: NADP; oxidoreductase

Query Match 41.1%; Score 23; DB 2; Length 10;
Best Local Similarity 40.0%; Pred. No. 4e+02;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```
Qy      1 IDNVKKARVQ 10
          :: | | :||
Db      1 VEKVSKKQVQ 10
```

RESULT 2
S70251
nitrogenase (EC 1.18.6.1) iron protein - Anabaena variabilis (fragment)
C;Species: Anabaena variabilis
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 19-Jan-2001
C;Accession: S70251
R;Schrautemeier, B.; Neveling, U.; Schmitz, S.
Mol. Microbiol. 18, 357-369, 1995
A;Title: Distinct and differently regulated Mo-dependent nitrogen-fixing systems evol egions as part of the nif1/2 gene clusters.
A;Reference number: S70242; MUID:96296457
A;Accession: S70251
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-10 <SCH>
A;Cross-references: EMBL:U25160; NID:g1160347; PIDN:AAC43540.1; PID:g1160348
A;Experimental source: ATCC 29413
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1995
C;Genetics:
A;Gene: nifH2
C;Keywords: ATP; iron-sulfur protein; nitrogen fixation; oxidoreductase

Query Match 39.3%; Score 22; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 6.1e+02;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 IDNVKK 6
```

```
          :|  | ||
Db        4 DVPTAAVQ 11


RESULT  14
A60757
enterotoxin C-1 - Staphylococcus aureus (fragments)
C;Species: Staphylococcus aureus
C;Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 30-Sep-1993
C;Accession: A60757
R;Bohach, G.A.; Handley, J.P.; Schlievert, P.M.
Infect. Immun. 57, 23-28, 1989
A;Title: Biological and immunological properties of the carboxyl terminus of staphylococ
A;Reference number: A60757; MUID:89079292
A;Accession: A60757
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-12 <BOH>


  Query Match             26.8%;  Score 15;  DB 2;  Length 12;
  Best Local Similarity   75.0%;  Pred. No. 1.5e+04;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        2 DNVK 5
            | ||
Db        3 DKVK 6


RESULT  15
S18722
matK protein - beechdrops plastid (fragment)
C;Species: plastid Epifagus virginiana (beechdrops)
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 20-Feb-1995
C;Accession: S18722
R;Morden, C.W.; Wolfe, K.H.; dePamphilis, C.W.; Palmer, J.D.
EMBO J. 10, 3281-3288, 1991
A;Title: Plastid translation and transcription genes in a non-photosynthetic plant: inta
A;Reference number: S17794; MUID:92007779
A;Accession: S18722
A;Molecule type: DNA
A;Residues: 1-12 <MOR>
A;Cross-references: EMBL:X61798
C;Genetics:
A;Gene: matK
A;Genome: plastid
C;Keywords: plastid


  Query Match             26.8%;  Score 15;  DB 2;  Length 12;
  Best Local Similarity   40.0%;  Pred. No. 1.5e+04;
  Matches    2;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        1 IDNVK 5
            :| :|
Db        1 MDKIK 5


RESULT  16
S65626
phosphorylase b - rabbit (fragments)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 13-Mar-1997
C;Accession: S65626
R;Luo, S.; Martin, B.L.; Senshu, T.; Graves, D.J.
Arch. Biochem. Biophys. 318, 362-369, 1995
A;Title: Enzymatic deimination of glycogen phosphorylase and a peptide of the phosphoryl
A;Reference number: S65626; MUID:95251385
A;Accession: S65626
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-3;4-12 <LUO>


  Query Match             26.8%;  Score 15;  DB 2;  Length 12;
  Best Local Similarity   50.0%;  Pred. No. 1.5e+04;
  Matches    2;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 IDNV 4
            ::||
Db        8 VENV 11


RESULT  17
B60110
repetitive protein antigen 61 - Trypanosoma cruzi (fragment)
C;Species: Trypanosoma cruzi
C;Date: 10-Nov-1992 #sequence_revision 10-Nov-1992 #text_change 18-Jun-1993
C;Accession: B60110
R;Hoft, D.F.; Kim, K.S.; Otsu, K.; Moser, D.R.; Yost, W.J.; Blumin, J.H.; Donelson, J
Infect. Immun. 57, 1959-1967, 1989
A;Title: Trypanosoma cruzi expresses diverse repetitive protein antigens.
A;Reference number: A60110; MUID:89277508
A;Accession: B60110
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-6 <HOF>
A;Note: this is an example of a five residue tandem repeat from this protein; the act
C;Keywords: tandem repeat


  Query Match             25.0%;  Score 14;  DB 2;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        5 KKA 7
            |||
Db        4 KKA 6


RESULT  18
S19630
ribosomal protein L30 - Streptomyces griseus (fragment)
C;Species: Streptomyces griseus
C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 06-Jun-1997
C;Accession: S19630
R;Ochi, K.
Int. J. Syst. Bacteriol. 42, 144-150, 1992
A;Title: Electrophoretic heterogeneity of ribosomal protein AT-L30 among actinomycete
A;Reference number: S19630; MUID:92144363
A;Accession: S19630
A;Molecule type: protein
A;Residues: 1-7 <OCH>
A;Experimental source: strain IFO 13189
C;Superfamily: Escherichia coli ribosomal protein L30
C;Keywords: protein biosynthesis; ribosome


  Query Match             25.0%;  Score 14;  DB 2;  Length 7;
  Best Local Similarity   40.0%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        7 ARVQV 11
            ||:::
Db        1 ARLKI 5


RESULT  19
T10077
hypothetical protein N - Methylophilus methylotrophus (fragment)
C;Species: Methylophilus methylotrophus
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C;Accession: T10077
R;Chistoserdov, A.Y.; McIntire, W.S.; Mathews, F.S.; Lidstrom, M.E.
J. Bacteriol. 176, 4073-4080, 1994
A;Title: Organization of the methylamine utilization (mau) genes in Methylophilus met
```

A;Reference number: Z16936; MUID:94292427
A;Accession: T10077
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-8 <CHI>
A;Cross-references: EMBL:L26407; NID:g561931; PIDN:AAB46955.1; PID:g561933
A;Experimental source: strain W3A1
C;Genetics:
A;Gene: mauN

Query Match 25.0%; Score 14; DB 2; Length 8;
Best Local Similarity 40.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy      8 RVQVV 12
          |:| :
Db      3 RLQAI 7
```

RESULT 20
PS0253
glycine cleavage system protein H - rice (strain Nihonbare) (fragment)
N;Alternate names: glycine decarboxylase complex H protein
C;Species: Oryza sativa (rice)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 03-Feb-1994
C;Accession: PS0253
R;Tsugita, A.
submitted to JIPID, April 1993
A;Reference number: PS0206
A;Accession: PS0253
A;Molecule type: protein
A;Residues: 1-9 <TSU>

Query Match 25.0%; Score 14; DB 2; Length 9;
Best Local Similarity 40.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 IDNVK 5
          :| :|
Db      4 LDGLK 8
```

Search completed: July 8, 2002, 11:43:58
Job time: 159 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 8, 2002, 11:48:58 ; Search time 13.1 Seconds
(without alignments)
35.468 Million cell updates/sec

Title: US-09-461-061A-3
Perfect score: 56
Sequence: 1 IDNVKKARVQVV 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 463

Minimum DB seq length: 0
Maximum DB seq length: 12

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	19	33.9	11	1	ES1_RAT	P56571 rattus norv
2	17	30.4	12	1	ULAL_MOUSE	P99032 mus musculu
3	16	28.6	12	1	HCYB_MEGCR	Q10584 megathura c
4	15	26.8	7	1	CHOX_ALCSP	P16101 alcaligenes
5	15	26.8	12	1	RS19_TOBBP	Q56251 tomato big
6	14	25.0	9	1	ULAD_HUMAN	P31929 homo sapien
7	14	25.0	10	1	FARP_LOCMI	P38553 locusta mig
8	14	25.0	10	1	LCMS_LEUMA	P21144 leucophaea
9	14	25.0	10	1	NEMS_DROME	P41494 drosophila
10	14	25.0	11	1	CSI5_BACSU	P81095 bacillus su
11	13	23.2	7	1	UH11_RAT	P56576 rattus norv
12	13	23.2	9	1	IPYR_RHOVI	P82992 rhodopseudo
13	13	23.2	9	1	RS10_SERMA	O68936 serratia ma
14	13	23.2	10	1	COXQ_RABIT	P80336 oryctolagus
15	13	23.2	10	1	UXA6_CHLTR	P38007 chlamydia t
16	13	23.2	12	1	FARI_CALVO	P41869 calliphora
17	13	23.2	12	1	OPS3_DROVI	P17645 drosophila
18	13	23.2	12	1	RS19_CLYEP	Q46490 clover yell
19	13	23.2	12	1	UH03_RAT	P56572 rattus norv
20	12	21.4	7	1	FARB_CALVO	P41866 calliphora
21	12	21.4	9	1	LPCA_STAAU	P36884 staphylococ
22	12	21.4	10	1	ANG1_BOTJA	Q10581 bothrops ja
23	12	21.4	10	1	ANGT_BOVIN	P01017 bos taurus
24	12	21.4	10	1	ANGT_CHICK	P01018 gallus gall
25	12	21.4	10	1	COXA_ONCMY	P80328 oncorhynchu
26	12	21.4	10	1	FARC_CALVO	P41867 calliphora
27	12	21.4	10	1	FIBB_CERSI	P14537 ceratotheri
28	12	21.4	10	1	MALE_KLEPN	Q05564 klebsiella
29	12	21.4	10	1	NS1_MYCTU	P81135 mycobacteri
30	12	21.4	10	1	ODP2_BOVIN	P11180 bos taurus
31	12	21.4	10	1	RRPL_PHODV	P35946 phocine dis
32	12	21.4	10	1	TEMK_RANTE	P56923 rana tempor
33	12	21.4	10	1	UPA9_HUMAN	P30095 homo sapien
34	12	21.4	11	1	HS70_PINPS	P81672 pinus pinas
35	12	21.4	11	1	TKNA_GADMO	P28498 gadus morhu
36	12	21.4	12	1	DCML_PSECA	P19919 pseudomonas
37	12	21.4	12	1	LICH_BACLI	P82907 bacillus li
38	12	21.4	12	1	YZPY_ECOLI	P17776 escherichia
39	11	19.6	8	1	ACT_CARMA	P80709 carcinus ma
40	11	19.6	8	1	ANG2_BOTJA	Q10582 bothrops ja
41	11	19.6	8	1	RS1_ERWCH	P37985 erwinia chr
42	11	19.6	8	1	UPA1_HUMAN	P30087 homo sapien
43	11	19.6	9	1	CONO_CONGE	P05486 conus geogr
44	11	19.6	9	1	DCML_PSECF	P19913 pseudomonas
45	11	19.6	9	1	OXYA_SCYCA	P42996 scyliorhinu

ALIGNMENTS

RESULT 1
ES1_RAT
ID ES1_RAT STANDARD; PRT; 11 AA.
AC P56571;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE ES1 protein, mitochondrial (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE.
RC STRAIN=WISTAR; TISSUE=Heart;
RA Li X.-P., Pleissner K.-P., Scheler C., Regitz-Zagrosek V., Salikov J.,
RA Jungblut P.R.;
RL Submitted (SEP-1998) to the SWISS-PROT data bank.
CC -!- SUBCELLULAR LOCATION: Mitochondrial (Potential).
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN
CC (SPOT P2) IS: 8.9, ITS MW IS: 25 kDa.
CC -!- SIMILARITY: BELONGS TO THE ES1 FAMILY.
KW Mitochondrion.
FT NON_TER 11 11
SQ SEQUENCE 11 AA; 1142 MW; D862272D32C72DC2 CRC64;

Query Match 33.9%; Score 19; DB 1; Length 11;
Best Local Similarity 57.1%; Pred. No. 8.5e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy      6 KARVQVV 12
          :||| :|
Db      1 RARVALV 7
```

RESULT 2
ULAL_MOUSE
ID ULAL_MOUSE STANDARD; PRT; 12 AA.
AC P99032;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Unknown protein from 2D-page of liver tissue (Spot 2D-0014M9)
DE (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RC TISSUE=Liver;
RA Sanchez J.-C., Rouge V., Frutiger S., Hughes G.J., Yan J.X.,
RA Hoogland C., Appel R.D., Binz P.-A., Hochstrasser D.F.,
RA Cowthorne M.;
RL Submitted (AUG-1998) to the SWISS-PROT data bank.

```
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 5.0, ITS MW IS: 11.7 kDa.
DR   SWISS-2DPAGE; P99032; MOUSE.
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1324 MW;  DD6468EE9F75BAB6 CRC64;


  Query Match             30.4%;  Score 17;  DB 1;  Length 12;
  Best Local Similarity   75.0%;  Pred. No. 2.2e+03;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 DNVK 5
          |||:
Db      2 DNVQ 5


RESULT   3
HCYB_MEGCR
ID   HCYB_MEGCR     STANDARD;      PRT;     12 AA.
AC   Q10584;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hemocyanin B chain (KLH-B) (Fragment).
OS   Megathura crenulata (Giant keyhole limpet).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Archaeogastropoda;
OC   Fissurellidae; Megathura.
OX   NCBI_TaxID=55429;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=96208935; PubMed=8829804;
RA   Swerdlow R.D., Ebert R.F., Lee P., Bonaventura C., Miller K.I.;
RT   "Keyhole limpet hemocyanin: structural and functional
RT   characterization of two different subunits and multimers.";
RL   Comp. Biochem. Physiol. 113B:537-548(1996).
CC   -!- FUNCTION: HEMOCYANINS ARE COPPER-CONTAINING OXYGEN CARRIERS
CC       OCCURRING FREELY DISSOLVED IN THE HEMOLYMPH OF MANY MOLLUSKS AND
CC       ARTHROPODS.
CC   -!- SUBUNIT: DIDECAMERS AND EXTENDED MULTIMERS.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- TISSUE SPECIFICITY: HEMOLYMPH.
CC   -!- BIOTECHNOLOGY: Potent immunogen used classically as a carrier
CC       protein for haptens and more recently in human vaccines and for
CC       immunotherapy of bladder cancer.
CC   -!- SIMILARITY: BELONGS TO THE TYROSINASE FAMILY. HEMOCYANIN
CC       SUBFAMILY.
DR   InterPro; IPR000896; Hemocyanin.
DR   PROSITE; PS00209; HEMOCYANIN_1; PARTIAL.
DR   PROSITE; PS00210; HEMOCYANIN_2; PARTIAL.
KW   Oxygen transport; Transport; Copper; Glycoprotein;
KW   Hemolymph.
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1345 MW;  CBFBEAA44A432412 CRC64;


  Query Match             28.6%;  Score 16;  DB 1;  Length 12;
  Best Local Similarity   33.3%;  Pred. No. 3.3e+03;
  Matches    3;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 IDNVKKARV 9
          :| | :  |
Db      1 VDTVVRKNV 9


RESULT   4
CHOX_ALCSP
ID   CHOX_ALCSP     STANDARD;      PRT;     7 AA.
AC   P16101;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-APR-1990 (Rel. 14, Last annotation update)
DE   Choline oxidase (EC 1.1.3.17) (Fragment).
OS   Alcaligenes sp.
OC   Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
OC   Alcaligenes.
OX   NCBI_TaxID=512;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=81006769; PubMed=6997283;
RA   Ohta-Fukuyama M., Miyake Y., Emi S., Yamano T.;
RT   "Identification and properties of the prosthetic group of choline
RT   oxidase from Alcaligenes sp.";
RL   J. Biochem. 88:197-203(1980).
CC   -!- CATALYTIC ACTIVITY: Choline + O(2) = betaine aldehyde + H(2)O(2).
DR   PIR; A15398; A15398.
KW   Oxidoreductase.
FT   NON_TER      7      7
SQ   SEQUENCE   7 AA;  839 MW;  7415B1E457644AC0 CRC64;


  Query Match             26.8%;  Score 15;  DB 1;  Length 7;
  Best Local Similarity   42.9%;  Pred. No. 1e+05;
  Matches    3;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 DNVKKAR 8
          ||   :|
Db      1 DNPNHSR 7


RESULT   5
RS19_TOBBP
ID   RS19_TOBBP     STANDARD;      PRT;     12 AA.
AC   Q56251;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   30S ribosomal protein S19 (Fragment).
GN   RPSS OR RPS19.
OS   Tomato big bud phytoplasma.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC   Acholeplasmataceae; Phytoplasma.
OX   NCBI_TaxID=35770;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94350802; PubMed=8071198;
RA   Gundersen D.E., Lee I.M., Rehner S.A., Davis R.E., Kingsbury D.T.;
RT   "Phylogeny of mycoplasmalike organisms (phytoplasmas): a basis for
RT   their classification.";
RL   J. Bacteriol. 176:5244-5254(1994).
CC   -!- FUNCTION: PROTEIN S19 FORMS A COMPLEX WITH S13 THAT BINDS STRONGLY
CC       TO THE 16S RIBOSOMAL RNA (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE S19P FAMILY OF RIBOSOMAL PROTEINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L27004; AAA83952.1; -.
DR   InterPro; IPR002222; Ribosomal_S19.
DR   PROSITE; PS00323; RIBOSOMAL_S19; PARTIAL.
KW   Ribosomal protein; rRNA-binding.
FT   NON_TER      1      1
SQ   SEQUENCE   12 AA;  1470 MW;  4CC88BE908333AA3 CRC64;


  Query Match             26.8%;  Score 15;  DB 1;  Length 12;
  Best Local Similarity   37.5%;  Pred. No. 5e+03;
  Matches    3;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 NVKKARVQ 10
          | |  ::|
Db      3 NKKDKKMQ 10
```

```
RESULT   6
ULAD_HUMAN
ID   ULAD_HUMAN     STANDARD;      PRT;     9 AA.
AC   P31929;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Unknown protein from 2D-page of liver tissue (Spot 106) (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Liver;
RX   MEDLINE=94147969; PubMed=8313870;
RA   Hughes G.J., Frutiger S., Paquet N., Pasquali C., Sanchez J.-C.,
RA   Tissot J.-D., Bairoch A., Appel R.D., Hochstrasser D.F.;
RT   "Human liver protein map: update 1993.";
RL   Electrophoresis 14:1216-1222(1993).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 6, ITS MW IS: 15 kDa.
DR   SWISS-2DPAGE; P31929; HUMAN.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1129 MW;  D02DFB41B6D33322 CRC64;


  Query Match             25.0%;  Score 14;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 VKK 6
          |||
Db      2 VKK 4


RESULT   7
FARP_LOCMI
ID   FARP_LOCMI     STANDARD;      PRT;     10 AA.
AC   P38553;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   SchistoFLRFamide (PDVDHFLRF-amide) (Cardioexcitatory neuropeptide).
OS   Locusta migratoria (Migratory locust), and
OS   Schistocerca gregaria (Desert locust).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC   Acridomorpha; Acridoidea; Acrididae; Locusta.
OX   NCBI_TaxID=7004, 7010;
RN   [1]
RP   SEQUENCE.
RC   SPECIES=L.migratoria; TISSUE=Brain;
RX   MEDLINE=93324430; PubMed=7687352;
RA   Schoofs L., Holman G.M., Paemen L., Veelaert D., Amelinckx M.,
RA   de Loof A.;
RT   "Isolation, identification, and synthesis of PDVDHFLRFamide
RT   (SchistoFLRFamide) in Locusta migratoria and its association with the
RT   male accessory glands, the salivary glands, the heart, and the
RT   oviduct.";
RL   Peptides 14:409-421(1993).
RN   [2]
RP   SEQUENCE.
RC   SPECIES=S.gregaria; TISSUE=Thoracic nervous system;
RX   MEDLINE=89246543; PubMed=2719702;
RA   Robb S., Packman L.C., Evans P.D.;
RT   "Isolation, primary structure and bioactivity of schistoflrf-amide, a
RT   FMRF-amide-like neuropeptide from the locust, Schistocerca
RT   gregaria.";
RL   Biochem. Biophys. Res. Commun. 160:850-856(1989).
CC   -!- FUNCTION: MUSCLE INHIBITING AGENT. INVOLVED IN THE NEURAL CONTROL
CC       OF THE VISCERAL MUSCLES OF THE HEART, ACCESSORY GLANDS AND
CC       OVIDUCT. MAY BE INVOLVED IN THE REGULATION OF SALIVA SECRETION.
CC   -!- TISSUE SPECIFICITY: FOUND IN AXONS OF THE MALE ACCESSORY GLANDS,
CC       THE SALIVARY GLANDS, THE HEART, AND THE OVIDUCT.
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
DR   PIR; A32543; A32543.
KW   Neuropeptide; Amidation.
FT   MOD_RES      10     10       AMIDATION.
SQ   SEQUENCE   10 AA;  1244 MW;  D3C51729D2C1EAB2 CRC64;


  Query Match             25.0%;  Score 14;  DB 1;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 6.3e+03;
  Matches    2;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 IDNV 4
          :|:|
Db      3 VDHV 6


RESULT   8
LCMS_LEUMA
ID   LCMS_LEUMA     STANDARD;      PRT;     10 AA.
AC   P21144; P41497;
DT   01-MAY-1991 (Rel. 18, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Leucomyosuppressin (LMS) (LeM-MS).
OS   Leucophaea maderae (Madeira cockroach).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC   Blaberoidea; Blaberidae; Leucophaea.
OX   NCBI_TaxID=6988;
RN   [1]
RP   SEQUENCE, AND SYNTHESIS.
RC   TISSUE=Head;
RA   Holman G.M., Cook B.J., Nachman R.J.;
RT   "Isolation, primary structure and synthesis of leucomyosuppressin,
RT   an insect neuropeptide that inhibits spontaneous contractions of the
RT   cockroach hindgut.";
RL   Comp. Biochem. Physiol. 85C:329-333(1986).
CC   -!- FUNCTION: INHIBITS THE SPONTANEOUS CONTRACTIONS OF COCKROACH
CC       PROTODEUM (HINDGUT).
KW   Neuropeptide; Amidation.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      10     10       AMIDATION.
SQ   SEQUENCE   10 AA;  1275 MW;  D3C45229D2C1EAB2 CRC64;


  Query Match             25.0%;  Score 14;  DB 1;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 6.3e+03;
  Matches    2;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 IDNV 4
          :|:|
Db      3 VDHV 6


RESULT   9
NEMS_DROME
ID   NEMS_DROME     STANDARD;      PRT;     10 AA.
AC   P41494;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Dromyosuppressin (Neomyosuppressin) (NEB-MS).
GN   DMS OR NEMS.
OS   Drosophila melanogaster (Fruit fly), and
OS   Sarcophaga bullata (Grey flesh fly) (Neobellieria bullata).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
```

```
OX   NCBI_TaxID=7227, 7385;
RN   [1]
RP   SEQUENCE.
RC   SPECIES=D.melanogaster;
RX   MEDLINE=93002195; PubMed=1390001;
RA   Nichols R.;
RT   "Isolation and structural characterization of Drosophila
RT   TDVDHVFLRFamide and FMRFamide-containing neural peptides.";
RL   J. Mol. Neurosci. 3:213-218(1992).
RN   [2]
RP   SEQUENCE.
RC   SPECIES=S.bullata; TISSUE=Head;
RX   MEDLINE=93047886; PubMed=1358537;
RA   Fonagy A., Schoofs L., Proost P., Van Damme J., Bueds H., De Loof A.;
RT   "Isolation, primary structure and synthesis of neomyosuppressin, a
RT   myoinhibiting neuropeptide from the grey fleshfly, Neobellieria
RT   bullata.";
RL   Comp. Biochem. Physiol. 102C:239-245(1992).
CC   -!- FUNCTION: MYOINHIBITING NEUROPEPTIDE.
DR   FlyBase; FBgn0011581; Dms.
KW   Neuropeptide; Amidation.
FT   MOD_RES      10     10       AMIDATION.
SQ   SEQUENCE   10 AA;  1248 MW;  D3C00329D2C1EAB2 CRC64;


  Query Match             25.0%;  Score 14;  DB 1;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 6.3e+03;
  Matches    2;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 IDNV 4
          :|:|
Db      3 VDHV 6


RESULT  10
CSI5_BACSU
ID   CSI5_BACSU     STANDARD;      PRT;     11 AA.
AC   P81095;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Cold shock protein CSI5 (11 kDa cold shock protein) (Fragment).
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=168 / JH642;
RA   Graumann P.L., Schmid R., Marahiel M.A.;
RL   Submitted (OCT-1997) to the SWISS-PROT data bank.
RN   [2]
RP   CHARACTERIZATION.
RC   STRAIN=168 / JH642;
RX   MEDLINE=96345629; PubMed=8755892;
RA   Graumann P., Schroeder K., Schmid R., Marahiel M.A.;
RT   "Cold shock stress-induced proteins in Bacillus subtilis.";
RL   J. Bacteriol. 178:4611-4619(1996).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- INDUCTION: IN RESPONSE TO LOW TEMPERATURE.
DR   SubtiList; BG?????; ????.
KW   Complete proteome.
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  1360 MW;  15F6ECEE6322C330 CRC64;


  Query Match             25.0%;  Score 14;  DB 1;  Length 11;
  Best Local Similarity   66.7%;  Pred. No. 6.9e+03;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 NVK 5
          |:|
Db      3 NIK 5


RESULT  11
UH11_RAT
ID   UH11_RAT       STANDARD;      PRT;     7 AA.
AC   P56576;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Unknown protein from 2D-page of heart tissue (Spot P11) (Fragment).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=WISTAR; TISSUE=Heart;
RA   Li X.-P., Pleissner K.-P., Scheler C., Regitz-Zagrosek V., Salikov J.,
RA   Jungblut P.R.;
RL   Submitted (SEP-1998) to the SWISS-PROT data bank.
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 8.5, ITS MW IS: 42 kDa.
FT   UNSURE        2      2       OR A.
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  775 MW;  6866DB040DC5A6B0 CRC64;


  Query Match             23.2%;  Score 13;  DB 1;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 1e+05;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      5 KKARVQ 10
          : || |
Db      1 QSAREQ 6


RESULT  12
IPYR_RHOVI
ID   IPYR_RHOVI     STANDARD;      PRT;     9 AA.
AC   P82992;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Inorganic pyrophosphatase (EC 3.6.1.1) (Pyrophosphate phospho-
DE   hydrolase) (PPase) (Fragment).
GN   PPA.
OS   Rhodopseudomonas viridis.
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Hyphomicrobium group; Blastochloris.
OX   NCBI_TaxID=1079;
RN   [1]
RP   SEQUENCE, FUNCTION, AND SUBCELLULAR LOCATION.
RA   Gomez R., Losada M., Serrano A.;
RL   Submitted (JUN-2001) to the SWISS-PROT data bank.
CC   -!- CATALYTIC ACTIVITY: Diphosphate + H(2)O = 2 phosphate.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED MW OF THIS PROTEIN IS:
CC       23 kDa.
CC   -!- SIMILARITY: BELONGS TO THE PPASE FAMILY.
DR   InterPro; IPR001596; Pyrophosphatase.
DR   PROSITE; PS00387; PPASE; PARTIAL.
KW   Hydrolase.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1014 MW;  68F3EAA05DDAA044 CRC64;


  Query Match             23.2%;  Score 13;  DB 1;  Length 9;
  Best Local Similarity   42.9%;  Pred. No. 1e+05;
  Matches    3;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 IDNVKKA 7
          || :  |
Db      3 IDAIDXA 9
```

```
RESULT 13
RS10_SERMA
ID   RS10_SERMA     STANDARD;      PRT;     9 AA.
AC   O68936;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   30S ribosomal protein S10 (Fragment).
GN   RPSJ.
OS   Serratia marcescens.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Serratia.
OX   NCBI_TaxID=615;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Noorani S.M., Lindahl L., Zengel J.M.;
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Involved in the binding of tRNA to the ribosomes (By
CC       similarity).
CC   -!- SIMILARITY: BELONGS TO THE S10P FAMILY OF RIBOSOMAL PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF058451; AAC14294.1; -.
DR   InterPro; IPR001848; Ribosomal_S10.
DR   PROSITE; PS00361; RIBOSOMAL_S10; PARTIAL.
KW   Ribosomal protein.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1214 MW;  DE3944004416D456 CRC64;


  Query Match             23.2%;  Score 13;  DB 1;  Length 9;
  Best Local Similarity   12.5%;  Pred. No. 1e+05;
  Matches    1;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 VKKARVQV 11
           ::  |::: 
Db       1 MQNQRIRI 8


RESULT 14
COXQ_RABIT
ID   COXQ_RABIT     STANDARD;      PRT;     10 AA.
AC   P80336;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   Cytochrome c oxidase polypeptide VIII-liver/heart (EC 1.9.3.1)
DE   (Fragment).
GN   COX8H.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Heart, and Liver;
RA   Freund R., Kadenbach B.;
RL   Submitted (MAR-1994) to the SWISS-PROT data bank.
CC   -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC       CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC       MITOCHONDRIAL ELECTRON TRANSPORT.
CC   -!- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC       c + 2 H(2)O.
KW   Oxidoreductase; Mitochondrion.
FT   NON_TER       10     10
SQ   SEQUENCE   10 AA;  1027 MW;  2C325CB40DC76338 CRC64;


  Query Match             23.2%;  Score 13;  DB 1;  Length 10;
  Best Local Similarity   75.0%;  Pred. No. 9.5e+03;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 KKAR 8
           | ||
Db       4 KPAR 7


RESULT 15
UXA6_CHLTR
ID   UXA6_CHLTR     STANDARD;      PRT;     10 AA.
AC   P38007;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Unknown protein from 2D-page from elementary body (Fragment).
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=813;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=L2/434/BU;
RA   Bini L., Santucci A., Magi B., Marzocchi B., Sanchez-Campillo M.,
RA   Comanducci M., Christianen G., Birkelund S., Vtretou E., Ratti G.,
RA   Pallini V.;
RL   Submitted (SEP-1994) to the SWISS-PROT data bank.
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 6.44, ITS MW IS: 38.6 kDa.
DR   Siena-2DPAGE; P38007; -.
FT   NON_TER       10     10
SQ   SEQUENCE   10 AA;  1243 MW;  DAD39A33304B5339 CRC64;


  Query Match             23.2%;  Score 13;  DB 1;  Length 10;
  Best Local Similarity   66.7%;  Pred. No. 9.5e+03;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 VKK 6
           :||
Db       6 IKK 8


RESULT 16
FARI_CALVO
ID   FARI_CALVO     STANDARD;      PRT;     12 AA.
AC   P41869;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   CalliMIRFamide 1.
OS   Calliphora vomitoria (Blue blowfly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Oestroidea; Calliphoridae; Calliphora.
OX   NCBI_TaxID=27454;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Thoracic ganglion;
RX   MEDLINE=92196111; PubMed=1549595;
RA   Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,
RA   Rehfeld J.F., Thorpe A.;
RT   "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
RT   neuropeptides (designated calliFMRFamides) from the blowfly
RT   Calliphora vomitoria.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
DR   PIR; E44787; E44787.
```

```
KW   Neuropeptide; Amidation.
FT   MOD_RES      12     12       AMIDATION.
SQ   SEQUENCE   12 AA;  1390 MW;  2DDDB316914AB5A7 CRC64;


  Query Match             23.2%;  Score 13;  DB 1;  Length 12;
  Best Local Similarity   66.7%;  Pred. No. 1.1e+04;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 DNV 4
          ||:
Db      7 DNM 9


RESULT  17
OPS3_DROVI
ID   OPS3_DROVI     STANDARD;      PRT;    12 AA.
AC   P17645;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   Opsin Rh3 (Inner R7 photoreceptor cells opsin) (Fragment).
GN   RH3.
OS   Drosophila virilis (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7244;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90249748; PubMed=2140105;
RA   Fortini M.E., Rubin G.M.;
RT   "Analysis of cis-acting requirements of the Rh3 and Rh4 genes reveals
RT   a bipartite organization to rhodopsin promoters in Drosophila
RT   melanogaster.";
RL   Genes Dev. 4:444-463(1990).
CC   -!- FUNCTION: VISUAL PIGMENTS ARE THE LIGHT-ABSORBING MOLECULES THAT
CC       MEDIATE VISION. THEY CONSIST OF AN APOPROTEIN, OPSIN, COVALENTLY
CC       LINKED TO CIS-RETINAL.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- MISCELLANEOUS: EACH DROSOPHILA EYE IS COMPOSED OF 800 FACETS OR
CC       OMMATIDIA. EACH OMMATIDIUM CONTAINS 8 PHOTORECEPTOR CELLS (R1-R8),
CC       THE R1 TO R6 CELLS ARE OUTER CELLS, WHILE R7 AND R8 ARE INNER
CC       CELLS.
CC   -!- MISCELLANEOUS: OPSIN RH3 IS SENSITIVE TO UV LIGHT.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC       OPSIN SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X51350; CAA35742.1; -.
DR   GCRDb; GCR_0779; -.
DR   FlyBase; FBgn0013091; Dvir\Rh3.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001760; Opsin.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; PARTIAL.
DR   PROSITE; PS00238; OPSIN; PARTIAL.
KW   Photoreceptor; Retinal protein; Transmembrane; Phosphorylation;
KW   Glycoprotein; G-protein coupled receptor; Vision.
FT   CARBOHYD     10     10       N-LINKED (GLCNAC. . .) (PROBABLE).
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1253 MW;  04024E43495865B0 CRC64;


  Query Match             23.2%;  Score 13;  DB 1;  Length 12;
  Best Local Similarity   75.0%;  Pred. No. 1.1e+04;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 IDNV 4
          | ||
Db      8 IGNV 11


RESULT  18
RS19_CLYEP
ID   RS19_CLYEP     STANDARD;      PRT;    12 AA.
AC   Q46490;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   30S ribosomal protein S19 (Fragment).
GN   RPSS OR RPS19.
OS   Clover yellow edge phytoplasma.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC   Acholeplasmataceae; Phytoplasma.
OX   NCBI_TaxID=35775;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94350802; PubMed=8071198;
RA   Gundersen D.E., Lee I.M., Rehner S.A., Davis R.E., Kingsbury D.T.;
RT   "Phylogeny of mycoplasmalike organisms (phytoplasmas): a basis for
RT   their classification.";
RL   J. Bacteriol. 176:5244-5254(1994).
CC   -!- FUNCTION: PROTEIN S19 FORMS A COMPLEX WITH S13 THAT BINDS STRONGLY
CC       TO THE 16S RIBOSOMAL RNA (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE S19P FAMILY OF RIBOSOMAL PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L27019; AAA83940.1; -.
DR   InterPro; IPR002222; Ribosomal_S19.
DR   PROSITE; PS00323; RIBOSOMAL_S19; PARTIAL.
KW   Ribosomal protein; rRNA-binding.
FT   NON_TER      1      1
SQ   SEQUENCE   12 AA;  1409 MW;  4CC478EFC8333AA3 CRC64;


  Query Match             23.2%;  Score 13;  DB 1;  Length 12;
  Best Local Similarity   33.3%;  Pred. No. 1.1e+04;
  Matches    2;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      5 KKARVQ 10
          |  ::|
Db      5 KDKKIQ 10


RESULT  19
UH03_RAT
ID   UH03_RAT       STANDARD;      PRT;    12 AA.
AC   P56572;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Unknown protein from 2D-page of heart tissue (Spot P3) (Fragment).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=WISTAR; TISSUE=Heart;
RA   Li X.-P., Pleissner K.-P., Scheler C., Regitz-Zagrosek V., Salikov J.,
RA   Jungblut P.R.;
RL   Submitted (SEP-1998) to the SWISS-PROT data bank.
```

```
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 8.3, ITS MW IS: 28 kDa.
FT   UNSURE          2       2
FT   UNSURE          9       9
FT   NON_TER        12      12
SQ   SEQUENCE   12 AA;  1255 MW;  46F58D101DC33053 CRC64;


  Query Match             23.2%;  Score 13;  DB 1;  Length 12;
  Best Local Similarity   50.0%;  Pred. No. 1.1e+04;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       6 KARVQV 11
             | :| |
Db       5 KIKVAV 10


RESULT  20
FARB_CALVO
ID   FARB_CALVO     STANDARD;      PRT;     7 AA.
AC   P41866;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   CalliFMRFamide 11.
OS   Calliphora vomitoria (Blue blowfly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Oestroidea; Calliphoridae; Calliphora.
OX   NCBI_TaxID=27454;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Thoracic ganglion;
RX   MEDLINE=92196111; PubMed=1549595;
RA   Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,
RA   Rehfeld J.F., Thorpe A.;
RT   "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
RT   neuropeptides (designated calliFMRFamides) from the blowfly
RT   Calliphora vomitoria.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
DR   PIR; B44787; B44787.
KW   Neuropeptide; Amidation.
FT   MOD_RES         7       7       AMIDATION.
SQ   SEQUENCE   7 AA;  926 MW;  69D40699C44AB700 CRC64;


  Query Match             21.4%;  Score 12;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 DN 3
             ||
Db       2 DN 3


Search completed: July  8, 2002, 11:48:59
Job time: 330 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 8, 2002, 11:48:39 ; Search time 39.63 Seconds
(without alignments)
52.383 Million cell updates/sec

Title: US-09-461-061A-3
Perfect score: 56
Sequence: 1 IDNVKKARVQVV 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 1842

Minimum DB seq length: 0
Maximum DB seq length: 12

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	21	37.5	12	8	Q37791	Q37791 larix eurol
2	21	37.5	12	8	Q36668	Q36668 pinus sylve
3	20	35.7	11	2	Q9R446	Q9r446 neisseria g
4	20	35.7	11	2	Q44090	Q44090 acholeplasm
5	20	35.7	12	16	O25179	O25179 helicobacte
6	18	32.1	10	4	Q9H121	Q9h121 homo sapien
7	18	32.1	10	10	P82434	P82434 nicotiana t
8	17	30.4	8	2	Q9AGP4	Q9agp4 arthrobacte
9	17	30.4	11	12	Q83083	Q83083 leucania se
10	17	30.4	12	8	P92457	P92457 ephedra sp.
11	16	28.6	8	3	Q05403	Q05403 saccharomyc
12	16	28.6	9	9	Q9XJN0	Q9xjn0 bacteriopha
13	16	28.6	10	7	Q95HG0	Q95hg0 papio anubi
14	16	28.6	10	7	Q95HF9	Q95hf9 papio anubi
15	16	28.6	10	7	Q95HF7	Q95hf7 papio anubi
16	16	28.6	10	7	Q95HF6	Q95hf6 papio anubi
17	16	28.6	10	7	Q95HF5	Q95hf5 papio anubi
18	16	28.6	10	7	Q95HF4	Q95hf4 papio anubi
19	16	28.6	11	2	Q9R5P3	Q9r5p3 serratia ma
20	16	28.6	11	10	O82070	O82070 triticum ae
21	16	28.6	11	12	Q86565	Q86565 varicella-z
22	16	28.6	12	8	Q37790	Q37790 larix eurol
23	16	28.6	12	8	Q36669	Q36669 pinus sylve
24	16	28.6	12	8	O03815	O03815 abies alba
25	16	28.6	12	8	O03816	O03816 metasequoia
26	16	28.6	12	10	P94011	P94011 arabidopsis
27	15	26.8	7	8	P92421	P92421 psathyrosta
28	15	26.8	7	8	P92385	P92385 hordeum mar
29	15	26.8	7	8	P92372	P92372 haynaldia v
30	15	26.8	7	8	P92403	P92403 lophopyrum
31	15	26.8	7	8	P92425	P92425 pseudoroegn
32	15	26.8	7	8	P92387	P92387 henrardia p
33	15	26.8	7	8	P92427	P92427 peridictyon
34	15	26.8	7	8	P92390	P92390 heteranthel
35	15	26.8	7	8	P92226	P92226 crithopsis
36	15	26.8	7	8	P92214	P92214 amblyopyrum
37	15	26.8	7	8	P92430	P92430 aegilops ta
38	15	26.8	7	8	P92221	P92221 bromus iner
39	15	26.8	7	8	P92442	P92442 taeniatheru
40	15	26.8	7	8	P92381	P92381 hordeum bra
41	15	26.8	7	8	P92393	P92393 hordeum vul
42	15	26.8	7	8	P92218	P92218 australopyr
43	15	26.8	7	8	P92440	P92440 thinopyrum
44	15	26.8	7	8	P92210	P92210 agropyron c
45	15	26.8	7	12	Q9YIR0	Q9yir0 human adeno

ALIGNMENTS

RESULT 1
Q37791
ID Q37791 PRELIMINARY; PRT; 12 AA.
AC Q37791;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TrEMBLrel. 01, Last annotation update)
DE CHLOROPLAST SUBUNIT OF LIGHT INDEPENDENT PROTOCHLOROPHYLLIDE REDUCTASE
DE (FRAGMENT).
GN CHLB.
OS Larix eurolepis.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Larix.
OX NCBI_TaxID=49226;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COTYLEDONS;
RA Karpinska B., Karpinski S., Hilgren J.E.;
RL Curr. Genet. 0:0-0(0).
DR EMBL; X98686; CAA67244.1; -.
DR EMBL; X98681; CAA67239.1; -.
DR EMBL; X98679; CAA67237.1; -.
KW Chloroplast.
FT NON_TER 1 1
FT NON_TER 12 12
SQ SEQUENCE 12 AA; 1445 MW; 0AB0FC0CC2276724 CRC64;

Query Match 37.5%; Score 21; DB 8; Length 12;
Best Local Similarity 50.0%; Pred. No. 2.1e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 IDNVKKAR 8
          : |: |||
Db      3 LKNLPKAR 10
```

RESULT 2

```
Q36668
ID   Q36668      PRELIMINARY;      PRT;    12 AA.
AC   Q36668;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CHLOROPLAST SUBUNIT OF LIGHT INDEPENDENT PROTOCHLOROPHYLLIDE REDUCTASE
DE   (FRAGMENT).
GN   CHLB.
OS   Pinus sylvestris (Scots pine).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX   NCBI_TaxID=3349;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=COTYLEDONS;
RX   MEDLINE=97263785; PubMed=9108142;
RA   Karpinska B., Karpinski S., Hallgren J.E.;
RT   "The chlB gene encoding a subunit of light-independent
RT   protochlorophyllide reductase is edited in chloroplast of conifers.";
RL   Curr. Genet. 31:343-347(1997).
DR   EMBL; X98683; CAA67240.1; -.
KW   Chloroplast.
FT   NON_TER       1      1
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1445 MW;  0AB0FC0CC2276724 CRC64;


  Query Match             37.5%;  Score 21;  DB 8;  Length 12;
  Best Local Similarity   50.0%;  Pred. No. 2.1e+03;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 IDNVKKAR 8
           : |: |||
Db       3 LKNLPKAR 10


RESULT   3
Q9R446
ID   Q9R446      PRELIMINARY;      PRT;    11 AA.
AC   Q9R446;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   CARBAMOYL-PHOSPHATE SYNTHASE SUBUNIT A (FRAGMENT).
GN   CARA.
OS   Neisseria gonorrhoeae.
OC   Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX   NCBI_TaxID=485;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MS11, AND FA1090;
RX   MEDLINE=95291461; PubMed=7773412;
RA   Lawson F.S., Billowes F.M., Dillon J.A.;
RT   "Organization of carbamoyl-phosphate synthase genes in Neisseria
RT   gonorrhoeae includes a large, variable intergenic sequence which is
RT   also present in other Neisseria species.";
RL   Microbiology 141:0-0(0).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MS11, AND FA1090;
RA   Brinkman F.S.L., Francis F.M., Dillon J.R.;
RT   "Complexity of the variable sequence between the carbamoyl-phosphate
RT   synthase genes of Neisseria species.";
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF029363; AAC78453.1; -.
DR   EMBL; AF029362; AAC78452.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   11 AA;  1178 MW;  0C07A8E3DDD33694 CRC64;


  Query Match             35.7%;  Score 20;  DB 2;  Length 11;
  Best Local Similarity   50.0%;  Pred. No. 3e+03;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 IDNVKKAR 8
           | |:| |:
Db       2 IGNMKAAK 9


RESULT   4
Q44090
ID   Q44090      PRELIMINARY;      PRT;    11 AA.
AC   Q44090;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL EXPORT SEGMENT (FRAGMENT).
OS   Acholeplasma laidlawii.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC   Acholeplasmataceae; Acholeplasma.
OX   NCBI_TaxID=2148;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A-EF22;
RA   Boyer M.J., Jarhede T.K., Tegman V., Wieslander A.;
RT   "Sequence regions from Acholeplasma laidlawii which restore export of
RT   beta-lactamase in Escherichia coli.";
RL   Submitted (JUN-1993) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z22875; CAA80495.1; -.
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  1234 MW;  5C9D2AE8A682C337 CRC64;


  Query Match             35.7%;  Score 20;  DB 2;  Length 11;
  Best Local Similarity   44.4%;  Pred. No. 3e+03;
  Matches    4;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       4 VKKARVQVV 12
           :|| :|  |
Db       1 MKKLKVMTV 9


RESULT   5
O25179
ID   O25179      PRELIMINARY;      PRT;    12 AA.
AC   O25179;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 1.4 KDA PROTEIN.
GN   HP0429.
OS   Helicobacter pylori (Campylobacter pylori).
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
OX   NCBI_TaxID=210;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=26695 / ATCC 700392;
RX   MEDLINE=97394467; PubMed=9252185;
RA   Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA   Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA   Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA   Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA   McKenney K., FitzGerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA   Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA   Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA   Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA   Venter J.C.;
RT   "The complete genome sequence of the gastric pathogen Helicobacter
RT   pylori.";
RL   Nature 388:539-547(1997).
DR   EMBL; AE000559; AAD07512.1; -.
DR   TIGR; HP0429; -.
KW   Hypothetical protein; Complete proteome.
```

```
SQ   SEQUENCE   12 AA;  1375 MW;  20959A84E2133338 CRC64;

  Query Match             35.7%;  Score 20;  DB 16;  Length 12;
  Best Local Similarity   40.0%;  Pred. No. 3.3e+03;
  Matches    4;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 DNVKKARVQV 11
          :| ||  :|:
Db      3 ENGKKEALQL 12


RESULT   6
Q9H121
ID   Q9H121      PRELIMINARY;      PRT;    10 AA.
AC   Q9H121;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   DJ309F20.1.2 (ISOFORM 2 OF GUANINE NUCLEOTIDE BINDING PROTEIN (G
DE   PROTEIN), ALPHA STIMULATING ACTIVITY POLYPEPTIDE 1) (FRAGMENT).
GN   GNAS1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Laird G.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL121917; CAC18783.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   10 AA;  1095 MW;  7809E3D322C7244B CRC64;

  Query Match             32.1%;  Score 18;  DB 4;  Length 10;
  Best Local Similarity   42.9%;  Pred. No. 6.3e+03;
  Matches    3;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 IDNVKKA 7
          ::|: ||
Db      2 LENLVKA 8


RESULT   7
P82434
ID   P82434      PRELIMINARY;      PRT;    10 AA.
AC   P82434;
DT   01-JUN-2000 (TrEMBLrel. 14, Created)
DT   01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   28 KDA CELL WALL PROTEIN (FRAGMENT).
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=CV. PETIT HAVANA;
RA   Blee K.A., Bonham V.A., Mitchell G.P., Robertson D., Slabas A.R.,
RA   Wojtaszek P., Bolwell G.P.;
RT   "Proteomic study of secondary cell wall proteins from transformed
RT   tobacco culture.";
RL   Planta 0:0-0(2000).
CC   -!- SUBCELLULAR LOCATION: CELL WALL.
CC   -!- TISSUE SPECIFICITY: XYLEM.
KW   Cell wall.
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1054 MW;  1A4B96DDDAB2C444 CRC64;

  Query Match             32.1%;  Score 18;  DB 10;  Length 10;
  Best Local Similarity   57.1%;  Pred. No. 6.3e+03;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 DNVKKAR 8
          :||  ||
Db      3 NNVDAAR 9


RESULT   8
Q9AGP4
ID   Q9AGP4      PRELIMINARY;      PRT;     8 AA.
AC   Q9AGP4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   SERINE HYDROXYMETHYLTRANSFERASE (FRAGMENT).
GN   GLYA.
OS   Arthrobacter sp. 1IN.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Micrococcineae; Micrococcaceae; Arthrobacter.
OX   NCBI_TaxID=153502;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1IN;
RA   Meskys R., Harris R.J., Casaite V., Basran J., Scrutton N.S.;
RT   "Genetic organization of the genes involved in dimethylglycine and
RT   sarcosine degradation in Arthrobacter spp.: implications for glycine
RT   betaine catabolism.";
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF329478; AAK16486.1; -.
KW   Transferase; Methyltransferase.
FT   NON_TER       1      1
SQ   SEQUENCE   8 AA;  898 MW;  6B18705333372457 CRC64;

  Query Match             30.4%;  Score 17;  DB 2;  Length 8;
  Best Local Similarity   75.0%;  Pred. No. 5.6e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 NVKK 6
          |:||
Db      2 NLKK 5


RESULT   9
Q83083
ID   Q83083      PRELIMINARY;      PRT;    11 AA.
AC   Q83083;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   P13 MINI PEPTIDE.
OS   Leucania separata nuclear polyhedrosis virus (LsNPV).
OC   Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC   Nucleopolyhedrovirus.
OX   NCBI_TaxID=41714;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96140622; PubMed=8572949;
RA   Wang J.W., Qi Y.P., Huang Y.X., Li S.D.;
RT   "Nucleotide sequence of a 1446 base pair SalI fragment and structure
RT   of a novel early gene of Leucania seperata nuclear polyhedrosis
RT   virus.";
RL   Arch. Virol. 140:2283-2291(1995).
DR   EMBL; U30303; AAA99737.1; -.
SQ   SEQUENCE   11 AA;  1339 MW;  F7BDBE0BD40DC401 CRC64;

  Query Match             30.4%;  Score 17;  DB 12;  Length 11;
  Best Local Similarity   57.1%;  Pred. No. 1.1e+04;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 NVKKARV 9
```

```
          |  :|||
Db      2 NRHRARV 8
```

```
RESULT  10
P92457
ID   P92457      PRELIMINARY;      PRT;    12 AA.
AC   P92457;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CHLOROPLAST SUBUNIT OF LIGHT-INDEPENDENT PROTOCHLOROPHYLLIDE REDUCTASE
DE   (FRAGMENT).
GN   CHLB GENE.
OS   Ephedra sp.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Gnetophyta; Gnetopsida; Ephedrales; Ephedraceae;
OC   Ephedra.
OX   NCBI_TaxID=41991;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97263785; PubMed=9108142;
RA   Karpinska B., Karpinski S., Hallgren J.E.;
RT   "The chlB gene encoding a subunit of light-independent
RT   protochlorophyllide reductase is edited in chloroplast of conifers.";
RL   Curr. Genet. 31:343-347(1997).
DR   EMBL; X98573; CAA67182.1; -.
KW   Chloroplast.
FT   NON_TER       1      1
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1441 MW;  164C1B7CC2276724 CRC64;
```

```
  Query Match             30.4%;  Score 17;  DB 8;  Length 12;
  Best Local Similarity   42.9%;  Pred. No. 1.2e+04;
  Matches    3;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 IDNVKKA 7
          : |: ||
Db      3 LQNLPKA 9
```

```
RESULT  11
Q05403
ID   Q05403      PRELIMINARY;      PRT;     8 AA.
AC   Q05403;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   DNA FOR ORF'S FROM CHROMOSOME XV (FRAGMENT).
GN   COQ3 AND YOL096C.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FY1679;
RX   MEDLINE=96021609; PubMed=8533473;
RA   Zumstein E., Pearson B.M., Kalogeropoulos A., Schweizer M.;
RT   "A 29.425 kb segment on the left arm of yeast chromosome XV contains
RT   more than twice as many unknown as known open reading frames.";
RL   Yeast 11:975-986(1995).
DR   EMBL; X83121; CAA58183.1; -.
DR   SGD; S0005456; COQ3.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  879 MW;  7B5322D2C441E058 CRC64;
```

```
  Query Match             28.6%;  Score 16;  DB 3;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 5.6e+05;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy      1 IDNVKK 6
          | || |
Db      2 IHNVVK 7
```

```
RESULT  12
Q9XJN0
ID   Q9XJN0      PRELIMINARY;      PRT;     9 AA.
AC   Q9XJN0;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   P10 (FRAGMENT).
OS   bacteriophage phi-10.
OC   Viruses; dsRNA viruses; Cystoviridae; Cystovirus.
OX   NCBI_TaxID=90889;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99350412; PubMed=10419946;
RA   Mindich L., Qiao X., Qiao J., Onodera S., Romantschuk M.,
RA   Hoogstraten D.;
RT   "Isolation of additional bacteriophages with genomes of segmented
RT   double-stranded RNA.";
RL   J. Bacteriol. 181:4505-4508(1999).
DR   EMBL; AF125675; AAD22555.1; -.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1058 MW;  880E376AA720544A CRC64;
```

```
  Query Match             28.6%;  Score 16;  DB 9;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 IDNV 4
          :||:
Db      1 MDNI 4
```

```
RESULT  13
Q95HG0
ID   Q95HG0      PRELIMINARY;      PRT;    10 AA.
AC   Q95HG0;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MHC CLASS I ANTIGEN (FRAGMENT).
GN   PAAN-AG.
OS   Papio anubis (Olive baboon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Papio.
OX   NCBI_TaxID=9555;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=HEART;
RA   Langat D.K., Morales P.J., Fazleabas A.T., Mwenda J.M., Hunt J.S.;
RT   "The baboon placenta contains unique tissue-specific messages derived
RT   from the class Ib major histocompatibility complex gene, Paan-AG.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY055032; AAL23590.1; -.
KW   MHC.
FT   NON_TER       1      1
SQ   SEQUENCE   10 AA;  1226 MW;  5C7EB7B6C1ADD443 CRC64;
```

```
  Query Match             28.6%;  Score 16;  DB 7;  Length 10;
  Best Local Similarity   60.0%;  Pred. No. 1.5e+04;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 NVKKA 7
          |:| |
Db      1 NMKNA 5
```

```
RESULT 14
Q95HF9
ID   Q95HF9      PRELIMINARY;      PRT;    10 AA.
AC   Q95HF9;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MHC CLASS I ANTIGEN (FRAGMENT).
GN   PAAN-AG.
OS   Papio anubis (Olive baboon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Papio.
OX   NCBI_TaxID=9555;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RA   Langat D.K., Morales P.J., Fazleabas A.T., Mwenda J.M., Hunt J.S.;
RT   "The baboon placenta contains unique tissue-specific messages derived
RT   from the class Ib major histocompatibility complex gene, Paan-AG.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY055033; AAL23591.1; -.
KW   MHC.
FT   NON_TER       1      1
SQ   SEQUENCE   10 AA;  1226 MW;  5C7EB7B6C1ADD443 CRC64;


  Query Match             28.6%;  Score 16;  DB 7;  Length 10;
  Best Local Similarity   60.0%;  Pred. No. 1.5e+04;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 NVKKA 7
          |:| |
Db      1 NMKNA 5


RESULT 15
Q95HF7
ID   Q95HF7      PRELIMINARY;      PRT;    10 AA.
AC   Q95HF7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MHC CLASS I ANTIGEN (FRAGMENT).
GN   PAAN-AG.
OS   Papio anubis (Olive baboon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Papio.
OX   NCBI_TaxID=9555;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LUNG;
RA   Langat D.K., Morales P.J., Fazleabas A.T., Mwenda J.M., Hunt J.S.;
RT   "The baboon placenta contains unique tissue-specific messages derived
RT   from the class Ib major histocompatibility complex gene, Paan-AG.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY055035; AAL23593.1; -.
KW   MHC.
FT   NON_TER       1      1
SQ   SEQUENCE   10 AA;  1226 MW;  5C7EB7B6C1ADD443 CRC64;


  Query Match             28.6%;  Score 16;  DB 7;  Length 10;
  Best Local Similarity   60.0%;  Pred. No. 1.5e+04;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 NVKKA 7
          |:| |
Db      1 NMKNA 5


RESULT 16
Q95HF6
ID   Q95HF6      PRELIMINARY;      PRT;    10 AA.
AC   Q95HF6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MHC CLASS I ANTIGEN (FRAGMENT).
GN   PAAN-AG.
OS   Papio anubis (Olive baboon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Papio.
OX   NCBI_TaxID=9555;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SKELETAL MUSCLE;
RA   Langat D.K., Morales P.J., Fazleabas A.T., Mwenda J.M., Hunt J.S.;
RT   "The baboon placenta contains unique tissue-specific messages derived
RT   from the class Ib major histocompatibility complex gene, Paan-AG.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY055036; AAL23594.1; -.
KW   MHC.
FT   NON_TER       1      1
SQ   SEQUENCE   10 AA;  1226 MW;  5C7EB7B6C1ADD443 CRC64;


  Query Match             28.6%;  Score 16;  DB 7;  Length 10;
  Best Local Similarity   60.0%;  Pred. No. 1.5e+04;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 NVKKA 7
          |:| |
Db      1 NMKNA 5


RESULT 17
Q95HF5
ID   Q95HF5      PRELIMINARY;      PRT;    10 AA.
AC   Q95HF5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MHC CLASS I ANTIGEN (FRAGMENT).
GN   PAAN-AG.
OS   Papio anubis (Olive baboon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Papio.
OX   NCBI_TaxID=9555;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SPLEEN;
RA   Langat D.K., Morales P.J., Fazleabas A.T., Mwenda J.M., Hunt J.S.;
RT   "The baboon placenta contains unique tissue-specific messages derived
RT   from the class Ib major histocompatibility complex gene, Paan-AG.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY055037; AAL23595.1; -.
KW   MHC.
FT   NON_TER       1      1
SQ   SEQUENCE   10 AA;  1226 MW;  5C7EB7B6C1ADD443 CRC64;


  Query Match             28.6%;  Score 16;  DB 7;  Length 10;
  Best Local Similarity   60.0%;  Pred. No. 1.5e+04;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 NVKKA 7
          |:| |
Db      1 NMKNA 5
```

```
RESULT 18
Q95HF4
ID   Q95HF4      PRELIMINARY;      PRT;    10 AA.
AC   Q95HF4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MHC CLASS I ANTIGEN (FRAGMENT).
GN   PAAN-AG.
OS   Papio anubis (Olive baboon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Papio.
OX   NCBI_TaxID=9555;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=THYMUS;
RA   Langat D.K., Morales P.J., Fazleabas A.T., Mwenda J.M., Hunt J.S.;
RT   "The baboon placenta contains unique tissue-specific messages derived
RT   from the class Ib major histocompatibility complex gene, Paan-AG.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY055038; AAL23596.1; -.
KW   MHC.
FT   NON_TER       1      1
SQ   SEQUENCE   10 AA;  1226 MW;  5C7EB7B6C1ADD443 CRC64;


  Query Match             28.6%;  Score 16;  DB 7;  Length 10;
  Best Local Similarity   60.0%;  Pred. No. 1.5e+04;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 NVKKA 7
          |:| |
Db      1 NMKNA 5


RESULT 19
Q9R5P3
ID   Q9R5P3      PRELIMINARY;      PRT;    11 AA.
AC   Q9R5P3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   SM2=NUCLEASE.
OS   Serratia marcescens.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Serratia.
OX   NCBI_TaxID=615;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92134331; PubMed=1663739;
RA   Bannikova G.E., Blagova E.V., Dementiev A.A., Morgunova E.Yu.,
RA   Mikchailov A.M., Shlyapnikov S.V., Varlamov V.P., Vainshtein B.K.;
RT   "Two isoforms of Serratia marcescens nuclease. Crystallization and
RT   preliminary X-ray investigation of the enzyme.";
RL   Biochem. Int. 24:813-822(1991).
SQ   SEQUENCE   11 AA;  1179 MW;  6DF18EE04AA045BB CRC64;


  Query Match             28.6%;  Score 16;  DB 2;  Length 11;
  Best Local Similarity  100.0%;  Pred. No. 1.6e+04;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 IDN 3
          |||
Db      6 IDN 8


RESULT 20
O82070
ID   O82070      PRELIMINARY;      PRT;    11 AA.
AC   O82070;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   RNA POLYMERASE (EC 2.7.7.6) (FRAGMENT).
OS   Triticum aestivum (Wheat).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Triticum.
OX   NCBI_TaxID=4565;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. CHINESE SPRING;
RA   Young D.A., Allen R., Harvey A.J., Lonsdale D.M.;
RT   "Characterization of a gene encoding a single-subunit RNA polymerase
RT   from maize which is alternatively spliced.";
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ005344; CAA06489.1; -.
KW   Transferase; Nucleotidyltransferase.
FT   NON_TER       1      1
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  1329 MW;  CD96344923240AB2 CRC64;


  Query Match             28.6%;  Score 16;  DB 10;  Length 11;
  Best Local Similarity   30.0%;  Pred. No. 1.6e+04;
  Matches    3;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 IDNVKKARVQ 10
          :| | :  |:
Db      1 VDEVDRKLVK 10


Search completed: July  8, 2002, 11:48:40
Job time: 336 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 8, 2002, 11:42:57 ; Search time 48.86 Seconds
(without alignments)
27.280 Million cell updates/sec

Title: US-09-461-061A-3
Perfect score: 56
Sequence: 1 IDNVKKARVQVV 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 158732

Minimum DB seq length: 0
Maximum DB seq length: 12

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	56	100.0	12	21	AAY95407	Anti-angiogenic pe
2	30	53.6	11	21	AAB26620	Partial sequence #
3	26	46.4	8	21	AAB08552	Peptide identified
4	26	46.4	12	19	AAW64637	Synthetic SEB pept
5	24	42.9	11	22	AAE12505	Membrane transitin
6	24	42.9	12	16	AAR82025	Stearoyl-ACP-desat
7	23	41.1	7	14	AAR43469	Ro/SSA epitope 313
8	23	41.1	8	22	AAE08179	Peptide #78 from h
9	23	41.1	8	22	AAE08180	Peptide #79 from h
10	23	41.1	8	22	AAE08181	Peptide #80 from h
11	23	41.1	9	20	AAY01110	CTGF peptide fragm
12	23	41.1	10	22	AAU07043	Human MHC molecule
13	23	41.1	11	15	AAR54725	ScFv-Ad5 fusion pr
14	23	41.1	12	20	AAY01108	HBGF-0.8-P2 peptid
15	23	41.1	12	20	AAW89521	Human latent trans
16	23	41.1	12	21	AAB15882	Human chemokine de
17	23	41.1	12	21	AAY78078	Human lactoferrin
18	22	39.3	8	14	AAR43470	Ro/SSA epitope 315
19	22	39.3	8	22	AAE08182	Peptide #81 from h
20	22	39.3	9	16	AAR75015	Deletion mutant de
21	22	39.3	10	19	AAW64638	Synthetic SEB pept
22	22	39.3	10	22	AAM43000	Mycoplasma genital
23	22	39.3	10	22	AAG88264	Saccharomyces cere
24	22	39.3	11	22	AAE05947	Basic aminooxyacet
25	22	39.3	12	16	AAR75020	Deletion mutant de
26	22	39.3	12	18	AAW35174	Lys/Val diastereom
27	22	39.3	12	18	AAW35175	Lys/Val diastereom
28	22	39.3	12	18	AAW35176	Lys/Val diastereom
29	22	39.3	12	19	AAW82890	Antipathogenic pep
30	22	39.3	12	19	AAW82891	Antipathogenic pep
31	22	39.3	12	19	AAW82892	Antipathogenic pep
32	22	39.3	12	21	AAB17425	Antipathogenic pep
33	22	39.3	12	21	AAB17427	Antipathogenic pep
34	22	39.3	12	21	AAY78083	Human lactoferrin
35	22	39.3	12	21	AAY78096	Human lactoferrin
36	21	37.5	6	9	AAP81903	A second minor try
37	21	37.5	7	22	AAM44081	H11 binding site c
38	21	37.5	7	22	AAM44086	H11 binding site c
39	21	37.5	7	22	AAM44092	H11 binding site c
40	21	37.5	7	22	AAM44172	H11 binding site c
41	21	37.5	7	22	AAM44177	H11 binding site c
42	21	37.5	7	22	AAM44675	H11 binding site c
43	21	37.5	7	22	AAM44678	H11 binding site c
44	21	37.5	7	22	AAM44685	H11 binding site c
45	21	37.5	7	22	AAM45852	H11 binding site c

ALIGNMENTS

RESULT 1
AAY95407
ID AAY95407 standard; Peptide; 12 AA.
XX
AC AAY95407;
XX
DT 25-SEP-2000 (first entry)
XX
DE Anti-angiogenic peptide C-terminal fragment.
XX
KW Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;
KW endothelial cell proliferation; apoptosis; cancer; ocular disorder;
KW rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic;
KW therapy; human; D3 peptide.
XX
OS Homo sapiens.
XX
PN WO200035407-A2.
XX
PD 22-JUN-2000.
XX
PF 02-DEC-1999; 99WO-US28465.
XX
PR 16-DEC-1998; 98US-0112427.
XX
PA (UTEM) UNIV TEMPLE.
PA (MCCR/) MCCRAE R K.
XX
PI McCrae RK;
XX
DR WPI; 2000-442247/38.
XX
PT Composition for inhibiting angiogenesis and endothelial cell
PT proliferation, inducing endothelial cell apoptosis and treating cancer,

PT rheumatoid arthritis, and ocular disorders comprises a kininogen domain
PT 3 analog -
XX
PS Claim 3; Page 25; 44pp; English.
XX
CC The present sequence is that of a C-terminal fragment of a novel
CC anti-angiogenic D3 peptide (see AAY95408) derived from human high
CC mol.wt. kininogen (HK) domain 3 (see AAY95426). The full-length D3
CC peptide inhibits endothelial cell proliferation and thus possesses
CC anti-angiogenic activity. It is an example of peptides of the
CC invention (see AAY95405-26) that are analogues of certain sites in
CC the HK domain 3. The peptides inhibit endothelial cell proliferation
CC and may also induce endothelial cell apoptosis. Compositions
CC including the peptides are used in claimed methods for inhibiting
CC angiogenesis, inhibiting endothelial cell proliferation, and
CC inducing endothelial cell apoptosis. Cancer, rheumatoid arthritis,
CC and ocular disorders characterized by undesired vascularization of
CC the retina are treated.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 56; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IDNVKKARVQVV 12
||||||||||||
Db 1 idnvkkarvqvv 12

RESULT 2
AAB26620
ID AAB26620 standard; Peptide; 11 AA.
XX
AC AAB26620;
XX
DT 22-JAN-2001 (first entry)
XX
DE Partial sequence #20 of Breast Cancer-Associated Protein Isoform.
XX
KW BPI; breast cancer-associated protein isoform; gene therapy;
KW breast cancer.
XX
OS Homo sapiens.
XX
PN WO200055628-A1.
XX
PD 21-SEP-2000.
XX
PF 13-MAR-2000; 2000WO-GB00908.
XX
PR 12-MAR-1999; 99GB-0005817.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Amess B, Townsend RR, Parekh RB, Waterfield MD, O'Hare MJ;
XX
DR WPI; 2000-602142/57.
XX
PT Screening, diagnosis of breast cancer and monitoring the effectiveness
PT of breast cancer therapy, involves detecting breast cancer-associated
PT features and breast cancer-associated protein isoforms -
XX
PS Disclosure; Page 13; 86pp; English.
XX
CC The present invention relates to the screening, diagnosis and prognosis
CC of breast cancer, for monitoring the effectiveness of breast cancer
CC treatment in a human, comprising identifying the presence or absence
CC of breast cancer-associated features (BF) or breast cancer-associated
CC protein isoforms (BPIs). Antibodies derived from BF and BPIs may be
CC useful for the treatment and screening of breast cancer, in particular
CC metastatic breast cancer. The present sequence is the partial
CC sequence of a BPI.
XX
SQ Sequence 11 AA;

Query Match 53.6%; Score 30; DB 21; Length 11;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IDNVKK 6
||||||
Db 6 idnvkk 11

RESULT 3
AAB08552
ID AAB08552 standard; Peptide; 8 AA.
XX
AC AAB08552;
XX
DT 20-DEC-2000 (first entry)
XX
DE Peptide identified from an origin of prepro-bradykinine.
XX
KW Precursor peptide; polypeptide hormone; peptide identification.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "hydrogen attached"
FT Modified-site 8
FT /note= "amidated residue"
XX
PN WO200050636-A1.
XX
PD 31-AUG-2000.
XX
PF 24-FEB-2000; 2000WO-FR00460.
XX
PR 25-FEB-1999; 99US-0257525.
XX
PA (SCRC) SCRAS SOC CONSEILS RECH & APPL SCI.
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Camara Ferrer YJA, Thurieau C, Martinez J, Berge G, Goze C;
XX
DR WPI; 2000-572101/53.
XX
PT Identifying peptide with selected function, useful particularly for
PT C-amidated hormones, by screening database for combination of nucleic
PT acid and amino acid sequences -
XX
PS Claim 16; Page 20; 40pp; French.
XX
CC The specification describes a method for identifying a peptide having
CC a particular function. The method comprises preparing a database of
CC polynucleotides and polypeptides of unknown functions, screening the
CC database for a combination of nucleotides or amino acids indicative of
CC the peptide with a particular function, and identifying polynucleotides
CC and proteins which contain the peptide. The method is used to identify
CC precursor peptides with an amidated C-terminus, especially polypeptide
CC hormones, for studying physiologically active substances. The present
CC sequence represents a peptide which was identified using the method of
CC the invention.
XX
SQ Sequence 8 AA;

Query Match 46.4%; Score 26; DB 21; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       7 ARVQVV 12
           ||||||
Db       2 arvqvv 7


RESULT   4
AAW64637
ID   AAW64637 standard; peptide; 12 AA.
XX
AC   AAW64637;
XX
DT   23-OCT-1998  (first entry)
XX
DE   Synthetic SEB peptide p12(151-161).
XX
KW   Enterotoxin B; SEB; pyrogenic exotoxin; PET; protective immunity;
KW   toxic shock; toxin-mediated activation; T-cell; antagonist; inhibitor;
KW   therapeutic; vaccine; food poisoning.
XX
OS   Synthetic.
OS   Staphylococcus aureus.
XX
PN   WO9829444-A1.
XX
PD   09-JUL-1998.
XX
PF   30-DEC-1997;   97WO-IL00438.
XX
PR   30-DEC-1996;   96IL-0119938.
XX
PA   (YISS ) YISSUM RES & DEV CO.
XX
PI   Arad G,  Kaempfer R;
XX
DR   WPI; 1998-388042/33.
XX
PT   New peptide(s) derived from pyrogenic exotoxin - useful for, e.g.
PT   antagonising toxin-mediated activation of T cells and prevention or
PT   treatment of toxic shock caused by exotoxin(s)
XX
PS   Claim 8; Page 41; 68pp; English.
XX
CC   AAW64636-W64657 are peptides homologous to the amino acid sequence of a
CC   fragment of a pyrogenic exotoxin (PET), and derivatives of the peptide
CC   capable of eliciting protective immunity against toxic shock induced by
CC   PET or by a mixture of PETs. Such peptides are also capable of
CC   antagonising toxin-mediated activation of T-cells, inhibiting expression
CC   of pyrogenic toxin (PT)-induced mRNA encoded by IL-2, IFN-gamma or
CC   TNF-beta genes. The peptides may be used to prepare therapeutics or
CC   vaccines for the treatment of prophylaxis of toxin-mediated activation
CC   of T cells and eliciting protective immunity against toxic shock induced
CC   by PETs. They can also be used for the treatment of harmful effects
CC   (especially food poisoning) and toxic shock caused by PET. Antiserum to
CC   the peptides can also be used for alleviating toxic shock induced by
CC   PET.
XX
SQ   Sequence   12 AA;


  Query Match             46.4%;  Score 26;  DB 19;  Length 12;
  Best Local Similarity   75.0%;  Pred. No. 1.6e+02;
  Matches    6;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 NVKKARVQ 10
           | ||| ||
Db       2 nkkkatvq 9


RESULT   5
AAE12505
ID   AAE12505 standard; peptide; 11 AA.
XX
AC   AAE12505;
XX
DT   03-JAN-2002  (first entry)
XX
DE   Membrane transiting antiviral peptide P22.
XX
KW   Membrane transiting peptide; virucide; antiviral; Herpes Simplex Virus;
KW   HSV; HIV; Human Immunodeficiency Virus; CMV; cytomegalovirus.
XX
OS   Unidentified.
XX
PN   WO200157072-A2.
XX
PD   09-AUG-2001.
XX
PF   06-FEB-2001; 2001WO-US03813.
XX
PR   07-FEB-2000; 2000US-180823P.
PR   22-FEB-2000; 2000US-184057P.
XX
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Brandt C,  Bultmann H;
XX
DR   WPI; 2001-638840/73.
XX
PT   Peptides comprising membrane transiting peptides useful for treating or
PT   preventing a virus infection, e.g., human immunodeficiency virus,
PT   herpes simplex virus and cytomegalovirus -
XX
PS   Claim 7; Page 7; 43pp; English.
XX
CC   The invention relates to peptides comprising membrane transiting
CC   peptides with antiviral properties. The peptides are useful for
CC   treating or preventing a virus infection in a warm blooded animal,
CC   e.g., enveloped viruses such as human immunodeficiency virus (HIV),
CC   herpes simplex virus (HSV), cytomegalovirus (CMV) and non-enveloped
CC   virus. Preferably, the peptides are useful for treating or preventing
CC   infections from one or more HSVs. The antiviral peptides are used for
CC   treating viral infections of the skin or part of the oral or genital
CC   cavity. The present sequence is membrane transiting antiviral peptide.
XX
SQ   Sequence   11 AA;


  Query Match             42.9%;  Score 24;  DB 22;  Length 11;
  Best Local Similarity   75.0%;  Pred. No. 3.5e+02;
  Matches    6;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 KKARVQVV 12
           ||| | ||
Db       3 kkaavavv 10


RESULT   6
AAR82025
ID   AAR82025 standard; Peptide; 12 AA.
XX
AC   AAR82025;
XX
DT   10-APR-1996  (first entry)
XX
DE   Stearoyl-ACP-desaturase N-terminal peptide (aa5-16).
XX
KW   Stearoyl-ACP-desaturase; soybean; seed oil; vegetable oil;
KW   fatty acid; stearic acid; transgenic plant; crop improvement..
XX
OS   Synthetic.
XX
PN   US5443974-A.
XX
PD   22-AUG-1995.
XX
PF   25-MAY-1990;   90US-0529049.
```

```
XX
PR   11-DEC-1992;   92US-0995657.
PR   25-MAY-1990;   90US-0529049.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Hitz WD,  Perez-Grau L,  Yadav NS;
XX
DR   WPI; 1995-302121/39.
XX
PT   New isolated soybean stearoyl-ACP desaturase gene - use for the
PT   prodn. of seed oil contg. altered levels of satd. and unsatd. fatty
PT   acids
XX
PS   Disclosure; Column 43-44; 25pp; English.
XX
CC   Amino acids 5-16 (AAR82025) of the soybean leaf stearoyl-ACP-
CC   desaturase mature enzyme were used to design a set of degenerate
CC   DNA probes (AAT00401).  The probes were utilized in the screening of
CC   soybean leaf cDNA expression libraries, leading to the isolation of
CC   a clone (AAT00395) coding for the precursor enzyme (AAR82024).
XX
SQ   Sequence   12 AA;


  Query Match             42.9%;  Score 24;  DB 16;  Length 12;
  Best Local Similarity   50.0%;  Pred. No. 3.8e+02;
  Matches    3;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 IDNVKK 6
            ::|:||
Db      3 venikk 8


RESULT   7
AAR43469
ID   AAR43469 standard; peptide; 7 AA.
XX
AC   AAR43469;
XX
DT   12-MAY-1994  (first entry)
XX
DE   Ro/SSA epitope 313.
XX
KW   Linear; epitope; 60 kD; Ro/SSA; La/SSB; autoantigen; E/F; G; 70 kD;
KW   nuclear ribonucleoprotein; nRNP; Sm B/B'; polypeptide; antigen; D;
KW   systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B; B';
KW   RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.
XX
OS   Homo sapiens.
XX
PN   WO9321223-A.
XX
PD   28-OCT-1993.
XX
PF   13-APR-1993;   93WO-US03484.
XX
PR   13-APR-1992;   92US-0867819.
XX
PA   (OKLA ) UNIV OKLAHOMA STATE.
XX
PI   Harley JB;
XX
DR   WPI; 1993-351658/44.
XX
PT   New linear epitope(s) for human auto-antibodies - from the
PT   Ro/SSA, La/SSB and Sm B/B' antigens and ribo:nucleoprotein, used
PT   for diagnosing and treating auto-immune disorders e.g. systemic
PT   lupus erythematosus
XX
PS   Claim 1; Page 31; 43pp; English.
XX
CC   The sequences given in AAR43391-562 are linear epitopes which are
CC   derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen,
CC   the 70 kD nuclear ribonucleoprotein (nRNP) and the Sm B/B'
CC   polypeptide.  These antigens are common in systemic lupus
CC   erythematosus (SLE) and closely related disorders.  The Ro/SSA family
CC   of proteins has been shown to have several molecular forms which are
CC   defined by the molecular weight of the antigen identified.  The major
CC   form has a molecular weight of 60 kD and two additional forms have
CC   molecular weights of 52 and 54 kD.  La/SSB is also a member of this
CC   group of autoantibodies and binds small RNAs with a polyuridine
CC   terminus.  La/SSB is bound by a third of the anti-Ro/SSA precipitin
CC   positive sera.  La/SSB has been shown to be a 46-50 kD monomeric
CC   phosphoprotein which associates with RNA polymerase III transcripts.
CC   Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and
CC   U5 RNA.  Anti-Sm antibodies may be directed against one or a
CC   combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD),
CC   E/F (11 kD doublet) and G (less than 10 kD).  These epitopes may be
CC   used for preventing, treating or screening autoimmune disorders,
CC   especially SLE or Sjogrens syndrome (SS).  They bind to a human
CC   autoantibody and may therefore be used as vaccines.
XX
SQ   Sequence   7 AA;


  Query Match             41.1%;  Score 23;  DB 14;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 6.4e+05;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 VKKARV 9
            :||||:
Db      2 lkkari 7


RESULT   8
AAE08179
ID   AAE08179 standard; peptide; 8 AA.
XX
AC   AAE08179;
XX
DT   01-NOV-2001  (first entry)
XX
DE   Peptide #78 from human ribonucleoprotein, 60 kD Ro/SSA.
XX
KW   Systemic lupus erythematosus; SLE; vaccine; immune response; therapeutic;
KW   autoimmune disorder; ribonucleoprotein; human; 60 kD Ro/SSA peptide.
XX
OS   Homo sapiens.
XX
PN   US6232522-B1.
XX
PD   15-MAY-2001.
XX
PF   30-NOV-1993;   93US-0160604.
XX
PR   31-JAN-1990;   90US-0472947.
PR   31-JAN-1991;   91US-0648205.
PR   13-APR-1992;   92US-0867819.
XX
PA   (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX
PI   Harley JB,  James JA,  Scofield RH;
XX
DR   WPI; 2001-335087/35.
XX
PT   Generating systemic lupus erythematosus animal model by immunizing
PT   non-human animal with non-immunoglobulin peptide having amino acid
PT   sequence of self-antigen bound by autoantibody population in early
PT   stage of disease -
XX
PS   Example 1; Column 16; 63pp; English.
XX
CC   The patent discloses a specific method of generating an animal model
CC   of systemic lupus erythematosus (SLE), comprising immunising a non-
CC   human animal with non-immunoglobulin peptide which comprises an
```

CC epitope immunoreactive with auto-antibody (AAb) from patient with
CC SLE. The epitope includes a region of self-antigen which is bound
CC by AAb population present in early stage in patient with SLE. The
CC method is used for generating an animal model of SLE. It is useful
CC for screening therapeutics effective in treating autoimmune disorders.
CC It is useful as a component in a diagnostic assay, as a therapeutic
CC (vaccine to block the AAbs produced, by eliciting immune response),
CC and in research on the possible causes of the autoimmune diseases.
CC The method is used to interrupt the course of an autoimmune response,
CC once autoimmunity against the autoantigen is established. The amino
CC acid sequences are used to make agents for neutralising circulating
CC antibodies or immobilised on substrates in extracorporeal devices for
CC specific removal of AAbs. The reagents identified by using the method
CC are useful in manufacturing and testing autoantigens. The method is
CC useful as models for screening of compounds which induce autoimmunity,
CC inhibit induction of autoimmunity, suppress autoimmunity. It is useful
CC in diagnosis of autoimmunity and as therapeutics for the treatment of
CC autoimmune disorders. The present sequence is a peptide from human ribo-
CC nucleoprotein, 60 kD Ro/SSA. This peptide is used in the exemplification
CC of the invention.
XX
SQ Sequence 8 AA;

Query Match 41.1%; Score 23; DB 22; Length 8;
Best Local Similarity 66.7%; Pred. No. 6.4e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 VKKARV 9
:||||:
Db 3 lkkari 8

RESULT 9
AAE08180
ID AAE08180 standard; peptide; 8 AA.
XX
AC AAE08180;
XX
DT 01-NOV-2001 (first entry)
XX
DE Peptide #79 from human ribonucleoprotein, 60 kD Ro/SSA.
XX
KW Systemic lupus erythematosus; SLE; vaccine; immune response; therapeutic;
KW autoimmune disorder; ribonucleoprotein; human; 60 kD Ro/SSA peptide.
XX
OS Homo sapiens.
XX
PN US6232522-B1.
XX
PD 15-MAY-2001.
XX
PF 30-NOV-1993; 93US-0160604.
XX
PR 31-JAN-1990; 90US-0472947.
PR 31-JAN-1991; 91US-0648205.
PR 13-APR-1992; 92US-0867819.
XX
PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX
PI Harley JB, James JA, Scofield RH;
XX
DR WPI; 2001-335087/35.
XX
PT Generating systemic lupus erythematosus animal model by immunizing
PT non-human animal with non-immunoglobulin peptide having amino acid
PT sequence of self-antigen bound by autoantibody population in early
PT stage of disease -
XX
PS Example 1; Column 16; 63pp; English.
XX
CC The patent discloses a specific method of generating an animal model
CC of systemic lupus erythematosus (SLE), comprising immunising a non-
CC human animal with non-immunoglobulin peptide which comprises an
CC epitope immunoreactive with auto-antibody (AAb) from patient with
CC SLE. The epitope includes a region of self-antigen which is bound
CC by AAb population present in early stage in patient with SLE. The
CC method is used for generating an animal model of SLE. It is useful
CC for screening therapeutics effective in treating autoimmune disorders.
CC It is useful as a component in a diagnostic assay, as a therapeutic
CC (vaccine to block the AAbs produced, by eliciting immune response),
CC and in research on the possible causes of the autoimmune diseases.
CC The method is used to interrupt the course of an autoimmune response,
CC once autoimmunity against the autoantigen is established. The amino
CC acid sequences are used to make agents for neutralising circulating
CC antibodies or immobilised on substrates in extracorporeal devices for
CC specific removal of AAbs. The reagents identified by using the method
CC are useful in manufacturing and testing autoantigens. The method is
CC useful as models for screening of compounds which induce autoimmunity,
CC inhibit induction of autoimmunity, suppress autoimmunity. It is useful
CC in diagnosis of autoimmunity and as therapeutics for the treatment of
CC autoimmune disorders. The present sequence is a peptide from human ribo-
CC nucleoprotein, 60 kD Ro/SSA. This peptide is used in the exemplification
CC of the invention.
XX
SQ Sequence 8 AA;

Query Match 41.1%; Score 23; DB 22; Length 8;
Best Local Similarity 66.7%; Pred. No. 6.4e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 VKKARV 9
:||||:
Db 2 lkkari 7

RESULT 10
AAE08181
ID AAE08181 standard; peptide; 8 AA.
XX
AC AAE08181;
XX
DT 01-NOV-2001 (first entry)
XX
DE Peptide #80 from human ribonucleoprotein, 60 kD Ro/SSA.
XX
KW Systemic lupus erythematosus; SLE; vaccine; immune response; therapeutic;
KW autoimmune disorder; ribonucleoprotein; human; 60 kD Ro/SSA peptide.
XX
OS Homo sapiens.
XX
PN US6232522-B1.
XX
PD 15-MAY-2001.
XX
PF 30-NOV-1993; 93US-0160604.
XX
PR 31-JAN-1990; 90US-0472947.
PR 31-JAN-1991; 91US-0648205.
PR 13-APR-1992; 92US-0867819.
XX
PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX
PI Harley JB, James JA, Scofield RH;
XX
DR WPI; 2001-335087/35.
XX
PT Generating systemic lupus erythematosus animal model by immunizing
PT non-human animal with non-immunoglobulin peptide having amino acid
PT sequence of self-antigen bound by autoantibody population in early
PT stage of disease -
XX
PS Example 1; Column 16; 63pp; English.
XX
CC The patent discloses a specific method of generating an animal model

CC of systemic lupus erythematosus (SLE), comprising immunising a non-
CC human animal with non-immunoglobulin peptide which comprises an
CC epitope immunoreactive with auto-antibody (AAb) from patient with
CC SLE. The epitope includes a region of self-antigen which is bound
CC by AAb population present in early stage in patient with SLE. The
CC method is used for generating an animal model of SLE. It is useful
CC for screening therapeutics effective in treating autoimmune disorders.
CC It is useful as a component in a diagnostic assay, as a therapeutic
CC (vaccine to block the AAbs produced, by eliciting immune response),
CC and in research on the possible causes of the autoimmune diseases.
CC The method is used to interrupt the course of an autoimmune response,
CC once autoimmunity against the autoantigen is established. The amino
CC acid sequences are used to make agents for neutralising circulating
CC antibodies or immobilised on substrates in extracorporeal devices for
CC specific removal of AAbs. The reagents identified by using the method
CC are useful in manufacturing and testing autoantigens. The method is
CC useful as models for screening of compounds which induce autoimmunity,
CC inhibit induction of autoimmunity, suppress autoimmunity. It is useful
CC in diagnosis of autoimmunity and as therapeutics for the treatment of
CC autoimmune disorders. The present sequence is a peptide from human ribo-
CC nucleoprotein, 60 kD Ro/SSA. This peptide is used in the exemplification
CC of the invention.
XX
SQ Sequence 8 AA;

Query Match 41.1%; Score 23; DB 22; Length 8;
Best Local Similarity 66.7%; Pred. No. 6.4e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 VKKARV 9
:||||:
Db 1 lkkari 6

RESULT 11
AAY01110
ID AAY01110 standard; peptide; 9 AA.
XX
AC AAY01110;
XX
DT 24-MAY-1999 (first entry)
XX
DE CTGF peptide fragment (residues 247-255).
XX
KW Heparin-binding growth factor; HBGF; connective tissue growth factor;
KW CTGF; pharmaceutical; wound healing; tissue formation; sclerotic; burn;
KW cell proliferative disorder; atherosclerosis; fibrotic; arthritis; IGF-1;
KW osteoporosis; skeletal disorder; scleroderma; liver cirrhosis; PDGF;
KW insulin-like growth factor; platelet-derived growth factor; TGF-beta;
KW EGF; epidermal growth factor; transforming growth factor beta; BFGF;
KW basic fibroblast growth factor; heparin; growth; placental membrane;
KW endometrial growth; pregnancy; endometrial-trophoblast interaction.
XX
OS Synthetic.
XX
PN WO9907407-A1.
XX
PD 18-FEB-1999.
XX
PF 06-AUG-1998; 98WO-US16423.
XX
PR 07-AUG-1997; 97US-0908526.
XX
PA (CHIL-) CHILDRENS HOSPITAL RES FOUND.
XX
PI Brigstock DA, Harding PA;
XX
DR WPI; 1999-167214/14.
XX
PT New substantially pure heparin-binding growth factor polypeptides -
PT useful for affecting wound healing and tissue formation, where
PT antagonists of the polypeptides are useful for, e.g. treating
PT sclerotic or cell proliferative disorders
XX
PS Example 5; Page 33; 53pp; English.
XX
CC The invention relates to a substantially pure heparin-binding growth
CC factor (HBGF) polypeptide (1 kDa) that corresponds to the C-terminal end
CC of a connective tissue growth factor (CTGF) protein. The HBGF
CC polypeptides form pharmaceutical compositions for affecting wound
CC healing and tissue formation. Antagonists of the polypeptides are useful
CC for treating sclerotic or cell proliferative disorders, atherosclerosis
CC or fibrotic conditions. Diseases and conditions modulated by HBGF
CC include arthritis, osteoporosis, and other skeletal disorders, burns.
CC Fibrotic conditions include scleroderma, arthritis and liver cirrhosis.
CC Cells treated are selected from epithelial (especially secretory),
CC muscle (especially smooth or cardiac muscle), connective tissue
CC (especially astroglia, fibroblast, osteoclast, osteoblast or chondrocyte
CC cells) and endothelial (especially endothelial) cells. Each treatment may
CC further comprise the use of growth factors selected from insulin-like
CC growth factor (IGF-1), platelet-derived growth factor (PDGF), epidermal
CC growth factor (EGF), transforming growth factor beta (TGF-beta) or basic
CC fibroblast growth factor (bFGF). The biological effect of HBGF can be
CC modulated using heparin at a concentration of 1-100 mg/ml. BBGF is used
CC for promoting endometrial growth and development of placental membranes
CC and promoting and maintaining pregnancy by facilitating endometrial-
CC trophoblast interaction. Sequences AAY01110-127 represent synthetic CTGF
CC peptide fragments.
XX
SQ Sequence 9 AA;

Query Match 41.1%; Score 23; DB 20; Length 9;
Best Local Similarity 42.9%; Pred. No. 6.4e+05;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 DNVKKAR 8
:|:|| :
Db 2 enikkgk 8

RESULT 12
AAU07043
ID AAU07043 standard; Peptide; 10 AA.
XX
AC AAU07043;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human MHC molecule HLA-B3501 binding 84P2A9 peptide #51.
XX
KW 84P2A9-related protein; prostate; testis; tissue; cancer; leukaemia;
KW tumour; kidney; brain; bone; skin; ovary; breast; pancreas; colon; lung;
KW cytostatic; gene therapy; antibody therapy; ribozyme; serum; blood;
KW single chain monoclonal antibody; urine.
XX
OS Homo sapiens.
XX
PN WO200155391-A2.
XX
PD 02-AUG-2001.
XX
PF 26-JAN-2001; 2001WO-US02651.
XX
PR 26-JAN-2000; 2000US-0178560.
XX
PA (UROG-) UROGENESYS INC.
XX
PI Jakobovits A, Afar DEH, Challita-eid PM, Levin E, Mitchell SC;
PI Hubert RS;
XX
DR WPI; 2001-502631/55.
XX
PT New 84P2A9 gene and its encoded protein, useful for diagnosing and
PT treating cancer, e.g. leukaemia and cancer of the prostate, testis,

```
PT   kidney, brain or bone, or for eliciting an immune response -
XX
PS   Example 12; Page 114; 149pp; English.
XX
CC   The polypeptide sequences represent the 84P2A9-related protein and
CC   peptide fragments of the protein. 84P2A9 exhibits prostate and testis
CC   specific expression in normal adult tissue, but it is also aberrantly
CC   expressed in many cancers including leukaemia and tumours of the
CC   prostate, testis, kidney, brain, bone, skin, ovary, breast, pancreas,
CC   colon and lung. The 84P2A9 polynucleotide, its related protein and
CC   peptide fragments and specific PCR primers are therefore useful for
CC   diagnosing and treating cancer. A vector comprising a polynucleotide
CC   which encodes a single chain monoclonal antibody, that immunospecifically
CC   binds to an 84P2A9-related protein, and a ribozyme capable of cleaving a
CC   polynucleotide having the 84P2A9 coding sequence, are both useful in the
CC   preparation of a composition for treating a patient with a cancer that
CC   expresses 84P2A9. The sequences can be used in diagnostic methods to
CC   monitor the level of 84P2A9 gene products in serum, blood, urine and
CC   tissue and to thereby detect the presence of cancerous cells.
XX
SQ   Sequence   10 AA;


  Query Match             41.1%;  Score 23;  DB 22;  Length 10;
  Best Local Similarity   33.3%;  Pred. No. 4.9e+02;
  Matches    3;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 VKKARVQVV 12
           ||| :::::
Db       2 vkkrklkii 10


RESULT  13
AAR54725
ID   AAR54725 standard; Protein; 11 AA.
XX
AC   AAR54725;
XX
DT   02-DEC-1994  (first entry)
XX
DE   ScFv-Ad5 fusion protein C.
XX
KW   Single chain Fv fragment; ScFv; Gene Therapy; Adenovirus;
KW   Penton fibre; Ad5; Fusion sites.
XX
OS   Adenovirus (Ad5).
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..6
FT                   /label= Ad5 fibre
FT   Peptide         7..10
FT                   /label= ScFv fragment
FT   Misc-difference 11
FT                   /label= X represents the remainder of the ScFv
FT                           fragment.
XX
PN   WO9410323-A.
XX
PD   11-MAY-1994.
XX
PF   04-NOV-1993;   93WO-GB02267.
XX
PR   04-NOV-1992;   92GB-0023084.
XX
PA   (IMCR ) IMPERIAL CANCER RES TECHNOLOGY.
XX
PI   Epenetos AA,  Spooner RA;
XX
DR   WPI; 1994-167477/20.
DR   N-PSDB; AAQ64764.
XX
PT   Virus with modified binding moiety specific for the target cells
PT   - used to deliver genes for gene therapy and cancer treatment
XX
PS   Example 1; Page 46; 110pp; English.
XX
CC   Displaying functional antibody fragments on the surface of
CC   recombinant retroviral particles could be used to target replication
CC   deficient virus to target cells for gene delivery.  The
CC   binding moiety (eg. a ScFv) is pref. external to the receptor for
CC   its host cell and fused to the virus direct or indirect by a spacer
CC   group. The binding site on the target cell is a cell specific
CC   antigen. Fusion C (AAQ64764 fused to AAQ64761) is at the end of the
CC   third repetitive unit of the shaft (co-ordinates 31323-4 (sequence
CC   co-ordinates taken from ADRCOMPGE_1) of the Ad5 fibre. The sequence
CC   between the PstI and XhoI sites is unique to the ScFv used. The first 6
CC   amino acid residues of the fusion A protein are from Ad5, the next
CC   4 from the ScFv fragment used.  X11 represents the remainder of the
CC   ScFv used.
XX
SQ   Sequence   11 AA;


  Query Match             41.1%;  Score 23;  DB 15;  Length 11;
  Best Local Similarity   50.0%;  Pred. No. 5.4e+02;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 VKKARVQV 11
           :|| :||:
Db       2 lkktkvql 9


RESULT  14
AAY01108
ID   AAY01108 standard; peptide; 12 AA.
XX
AC   AAY01108;
XX
DT   24-MAY-1999  (first entry)
XX
DE   HBGF-0.8-P2 peptide fragment obtained by N-terminal sequencing.
XX
KW   Heparin-binding growth factor; HBGF; connective tissue growth factor;
KW   CTGF; pharmaceutical; wound healing; tissue formation; sclerotic; burn;
KW   cell proliferative disorder; atherosclerosis; fibrotic; arthritis; IGF-1;
KW   osteoporosis; skeletal disorder; scleroderma; liver cirrhosis; PDGF;
KW   insulin-like growth factor; platelet-derived growth factor; TGF-beta;
KW   EGF; epidermal growth factor; transforming growth factor beta; BFGF;
KW   basic fibroblast growth factor; heparin; growth; placental membrane;
KW   endometrial growth; pregnancy; endometrial-trophoblast interaction.
XX
OS   Sus scrofa.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 9
FT                   /note= "unspecified"
XX
PN   WO9907407-A1.
XX
PD   18-FEB-1999.
XX
PF   06-AUG-1998;   98WO-US16423.
XX
PR   07-AUG-1997;   97US-0908526.
XX
PA   (CHIL-) CHILDRENS HOSPITAL RES FOUND.
XX
PI   Brigstock DA,  Harding PA;
XX
DR   WPI; 1999-167214/14.
XX
PT   New substantially pure heparin-binding growth factor polypeptides -
PT   useful for affecting wound healing and tissue formation, where
PT   antagonists of the polypeptides are useful for, e.g. treating
PT   sclerotic or cell proliferative disorders
XX
```

```
PS   Example 2; Page 27; 53pp; English.
XX
CC   The invention relates to a substantially pure heparin-binding growth
CC   factor (HBGF) polypeptide (1 kDa) that corresponds to the C-terminal end
CC   of a connective tissue growth factor (CTGF) protein. The HBGF
CC   polypeptides form pharmaceutical compositions for affecting wound healing
CC   and tissue formation. Antagonists of the polypeptides are useful for
CC   treating sclerotic or cell proliferative disorders, atherosclerosis or
CC   fibrotic conditions. Diseases and conditions modulated by HBGF include
CC   arthritis, osteoporosis, and other skeletal disorders, burns. Fibrotic
CC   conditions include scleroderma, arthritis and liver cirrhosis. Cells
CC   treated are selected from epithelial (especially secretory), muscle
CC   (especially smooth or cardiac muscle), connective tissue (especially
CC   astroglia, fibroblast, osteoclast, osteoblast or chondrocyte cells) and
CC   endothelial (especially endothelial) cells. Each treatment may further
CC   comprise the use of growth factors selected from insulin-like growth
CC   factor (IGF-1), platelet-derived growth factor (PDGF), epidermal growth
CC   factor (EGF), transforming growth factor beta (TGF-beta) or basic
CC   fibroblast growth factor (bFGF). The biological effect of HBGF can be
CC   modulated using heparin at a concentration of 1-100 mg/ml. BBGF is used
CC   for promoting endometrial growth and development of placental membranes
CC   and promoting and maintaining pregnancy by facilitating endometrial-
CC   trophoblast interaction. The present sequence represents a HBGF peptide
CC   fragment obtained by N-terminal sequencing.
XX
SQ   Sequence   12 AA;


  Query Match             41.1%;  Score 23;  DB 20;  Length 12;
  Best Local Similarity   42.9%;  Pred. No. 6e+02;
  Matches    3;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 DNVKKAR 8
           :|:|| :
Db       1 enikkgk 7


RESULT  15
AAW89521
ID   AAW89521 standard; peptide; 12 AA.
XX
AC   AAW89521;
XX
DT   11-MAR-1999  (first entry)
XX
DE   Human latent transforming growth factor-beta promoting peptide #6.
XX
KW   Human; latent transforming growth factor beta; TGF-beta; cancer;
KW   atherosclerosis; activity regulator; retinopathy; bone fracture;
KW   myocardial infarction; cerebral infarction; hepatic infarction;
KW   nephritis; HIV-associated kidney disease; rheumatoid arthritis.
XX
OS   Homo sapiens.
XX
PN   WO9851704-A1.
XX
PD   19-NOV-1998.
XX
PF   12-MAY-1998;   98WO-JP02089.
XX
PR   12-MAY-1997;   97JP-0120683.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO KK.
XX
PI   Sato Y,  Shibata K,  Yamasaki M;
XX
DR   WPI; 1999-059698/05.
XX
PT   Peptides promoting activation of latent transforming growth factor
PT   beta (TGF beta) - for screening potential drugs for treating
PT   TGF-associated diseases such as cancer and atherosclerosis
XX
PS   Claim 6; Page 30; 115pp; Japanese.
XX
CC   AAW89516 to AAW89531 represent peptides which promote the activation of
CC   latent transforming growth factor beta (TGF-beta). The peptides have
CC   the formula R1-A-R2 (I) where R1 = H or (optionally substituted)
CC   alkanoyl, arylcarbonyl, heteroarylcarbonyl, alkoxycarbonyl,
CC   aryloxycarbonyl or heteroaryloxycarbonyl; R2 = OH, or (optionally
CC   substituted) alkoxy or amino; A = a partial sequence from the precursor
CC   sequence of TGF-beta, in which up to 5 residues may be added, deleted
CC   or substituted, and side-chain amino or carboxy groups may be linked by
CC   an amide bond or side-chain mercapto groups may be linked by a
CC   disulphide bond. The peptides can be used for the screening of candidate
CC   drugs for the treatment and prevention of diseases with which TGF-beta
CC   is associated, such as cancer, diabetic complications (e.g.
CC   retinopathy), atherosclerosis, bone fractures, myocardial infarction,
CC   cerebral infarction, hepatic infarction, nephritis, HIV-associated kidney
CC   disease and rheumatoid arthritis.
XX
SQ   Sequence   12 AA;


  Query Match             41.1%;  Score 23;  DB 20;  Length 12;
  Best Local Similarity   33.3%;  Pred. No. 6e+02;
  Matches    3;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 VKKARVQVV 12
           ||: |:: :
Db       2 vkrkrieai 10


RESULT  16
AAB15882
ID   AAB15882 standard; Peptide; 12 AA.
XX
AC   AAB15882;
XX
DT   17-JAN-2001  (first entry)
XX
DE   Human chemokine derived peptide #34.
XX
KW   Macrophage recruitment; chemokine derivative; MCP-1; osteoporosis;
KW   monocyte chemoattractant protein-1; inflammation; atherosclerosis; HIV;
KW   AIDS; stroke; psoriasis; autoimmune disease; hypertension; endotoxaemia;
KW   basophil-mediated disease; myocardial infarction; acute ischaemia;
KW   rheumatoid arthritis; contraception.
XX
OS   Synthetic.
XX
PN   WO200042071-A2.
XX
PD   20-JUL-2000.
XX
PF   12-JAN-2000; 2000WO-US00821.
XX
PR   12-JAN-1999;   99US-0229071.
PR   17-MAR-1999;   99US-0271192.
PR   01-DEC-1999;   99US-0452406.
XX
PA   (NEOR-) NEORX CORP.
XX
PI   Grainger DJ,  Tatalick LM;
XX
DR   WPI; 2000-499101/44.
XX
PT   New peptide 3, amide and heterocyclic compounds and saccharide
PT   conjugates used for inhibiting chemokine induced activity and for
PT   treating e.g. stroke, vascular diseases, autoimmune diseases and tumour
PT   growth -
XX
PS   Disclosure; Fig 18; 387pp; English.
XX
CC   The present invention concerns the identification of a number of
CC   chemokines which can be used to produce derivatives, agonists and
CC   antagonists which are then useful in disease treatment. The chemokines
```

CC include sequences AAB15785-B15794, AAB15803-B15813 and AAB15831-B15848.
CC These chemokine derivatives can be used to treat diseases such as
CC autoimmune diseases, atherosclerosis, osteoporosis, HIV infection and
CC AIDS, psoriasis, inflammatory diseases, hypertension, basophil-mediated
CC diseases, endotoxaemia, myocardial infarction, acute ischaemia and
CC rheumatoid arthritis, and can be used to prevent strokes and as
CC contraceptives. The coding sequences for the chemokines can be used in
CC gene therapy for the same diseases, as well as in the production of
CC animal models.
XX
SQ Sequence 12 AA;

Query Match 41.1%; Score 23; DB 21; Length 12;
Best Local Similarity 33.3%; Pred. No. 6e+02;
Matches 3; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 4 VKKARVQVV 12
||: |:: :
Db 2 vkrkrieai 10

RESULT 17
AAY78078
ID AAY78078 standard; Peptide; 12 AA.
XX
AC AAY78078;
XX
DT 25-APR-2000 (first entry)
XX
DE Human lactoferrin derived peptide SEQ ID NO:78.
XX
KW Human; lactoferrin; modification; infection; inflammation; tumour;
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;
KW urinary tract infection; colitis; Candida infection; fungicidal;
KW bactericidal; preservative.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200001730-A1.
XX
PD 13-JAN-2000.
XX
PF 06-JUL-1999; 99WO-SE01230.
XX
PR 06-JUL-1998; 98SE-0002441.
PR 17-JUL-1998; 98SE-0002562.
PR 29-DEC-1998; 98SE-0004614.
XX
PA (ASCI-) A+ SCI INVEST AB.
XX
PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;
XX
DR WPI; 2000-147388/13.
XX
PT New peptides used for treatment and prevention of infections,
PT inflammations and tumors and for use in infant formula food -
XX
PS Claim 22; Page 35; 102pp; English.
XX
CC AAY78001 to AAY78100 represent peptides having sequences based on human
CC lactoferrin. The peptides are taken up in the intestine through
CC binding to specific lactoferrin receptors and are then transported
CC through the circulation. A medicinal product of the peptide or fragment
CC can be used for treating and/or prevention of infections (such as
CC urinary tract infections, colitis, and Candida infection on a mucosal
CC membrane), inflammations and/or tumours. The peptides can also be used
CC in food stuffs such as infant formula food. The peptides are also
CC fungicidal and bactericidal and may also be used as preservatives.
CC Even though native human lactoferrin have been shown to have desired
CC anti-inflammatory anti-infectious and anti-tumoural properties they
CC cannot be used clinically on a broad basis because of high production
CC costs. Therefore, provision of peptides based on lactoferrin would
CC enable them to be used for the same purposes as lactoferrin at lower
CC cost.
XX
SQ Sequence 12 AA;

Query Match 41.1%; Score 23; DB 21; Length 12;
Best Local Similarity 66.7%; Pred. No. 6e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 NVKKAR 8
|::|||
Db 7 nmrkar 12

RESULT 18
AAR43470
ID AAR43470 standard; peptide; 8 AA.
XX
AC AAR43470;
XX
DT 12-MAY-1994 (first entry)
XX
DE Ro/SSA epitope 315.
XX
KW Linear; epitope; 60 kD; Ro/SSA; La/SSB; autoantigen; E/F; G; 70 kD;
KW nuclear ribonucleoprotein; nRNP; Sm B/B'; polypeptide; antigen; D;
KW systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B; B';
KW RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.
XX
OS Homo sapiens.
XX
PN WO9321223-A.
XX
PD 28-OCT-1993.
XX
PF 13-APR-1993; 93WO-US03484.
XX
PR 13-APR-1992; 92US-0867819.
XX
PA (OKLA) UNIV OKLAHOMA STATE.
XX
PI Harley JB;
XX
DR WPI; 1993-351658/44.
XX
PT New linear epitope(s) for human auto-antibodies - from the
PT Ro/SSA, La/SSB and Sm B/B' antigens and ribo:nucleoprotein, used
PT for diagnosing and treating auto-immune disorders e.g. systemic
PT lupus erythematosus
XX
PS Claim 1; Page 31; 43pp; English.
XX
CC The sequences given in AAR43391-562 are linear epitopes which are
CC derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen,
CC the 70 kD nuclear ribonucleoprotein (nRNP) and the Sm B/B'
CC polypeptide. These antigens are common in systemic lupus
CC erythematosus (SLE) and closely related disorders. The Ro/SSA family
CC of proteins has been shown to have several molecular forms which are
CC defined by the molecular weight of the antigen identified. The major
CC form has a molecular weight of 60 kD and two additional forms have
CC molecular weights of 52 and 54 kD. La/SSB is also a member of this
CC group of autoantibodies and binds small RNAs with a polyuridine
CC terminus. La/SSB is bound by a third of the anti-Ro/SSA precipitin
CC positive sera. La/SSB has been shown to be a 46-50 kD monomeric
CC phosphoprotein which associates with RNA polymerase III transcripts.
CC Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and
CC U5 RNA. Anti-Sm antibodies may be directed against one or a
CC combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD),
CC E/F (11 kD doublet) and G (less than 10 kD). These epitopes may be
CC used for preventing, treating or screening autoimmune disorders,
CC especially SLE or Sjogrens syndrome (SS). They bind to a human

CC autoantibody and may therefore be used as vaccines.
XX
SQ Sequence 8 AA;

Query Match 39.3%; Score 22; DB 14; Length 8;
Best Local Similarity 80.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 KKARV 9
||||:
Db 1 kkari 5

RESULT 19
AAE08182
ID AAE08182 standard; peptide; 8 AA.
XX
AC AAE08182;
XX
DT 01-NOV-2001 (first entry)
XX
DE Peptide #81 from human ribonucleoprotein, 60 kD Ro/SSA.
XX
KW Systemic lupus erythematosus; SLE; vaccine; immune response; therapeutic;
KW autoimmune disorder; ribonucleoprotein; human; 60 kD Ro/SSA peptide.
XX
OS Homo sapiens.
XX
PN US6232522-B1.
XX
PD 15-MAY-2001.
XX
PF 30-NOV-1993; 93US-0160604.
XX
PR 31-JAN-1990; 90US-0472947.
PR 31-JAN-1991; 91US-0648205.
PR 13-APR-1992; 92US-0867819.
XX
PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX
PI Harley JB, James JA, Scofield RH;
XX
DR WPI; 2001-335087/35.
XX
PT Generating systemic lupus erythematosus animal model by immunizing
PT non-human animal with non-immunoglobulin peptide having amino acid
PT sequence of self-antigen bound by autoantibody population in early
PT stage of disease -
XX
PS Example 1; Column 16; 63pp; English.
XX
CC The patent discloses a specific method of generating an animal model
CC of systemic lupus erythematosus (SLE), comprising immunising a non-
CC human animal with non-immunoglobulin peptide which comprises an
CC epitope immunoreactive with auto-antibody (AAb) from patient with
CC SLE. The epitope includes a region of self-antigen which is bound
CC by AAb population present in early stage in patient with SLE. The
CC method is used for generating an animal model of SLE. It is useful
CC for screening therapeutics effective in treating autoimmune disorders.
CC It is useful as a component in a diagnostic assay, as a therapeutic
CC (vaccine to block the AAbs produced, by eliciting immune response),
CC and in research on the possible causes of the autoimmune diseases.
CC The method is used to interrupt the course of an autoimmune response,
CC once autoimmunity against the autoantigen is established. The amino
CC acid sequences are used to make agents for neutralising circulating
CC antibodies or immobilised on substrates in extracorporeal devices for
CC specific removal of AAbs. The reagents identified by using the method
CC are useful in manufacturing and testing autoantigens. The method is
CC useful as models for screening of compounds which induce autoimmunity,
CC inhibit induction of autoimmunity, suppress autoimmunity. It is useful
CC in diagnosis of autoimmunity and as therapeutics for the treatment of
CC autoimmune disorders. The present sequence is a peptide from human ribo-
CC nucleoprotein, 60 kD Ro/SSA. This peptide is used in the exemplification
CC of the invention.
XX
SQ Sequence 8 AA;

Query Match 39.3%; Score 22; DB 22; Length 8;
Best Local Similarity 80.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 KKARV 9
||||:
Db 1 kkari 5

RESULT 20
AAR75015
ID AAR75015 standard; peptide; 9 AA.
XX
AC AAR75015;
XX
DT 16-JAN-1996 (first entry)
XX
DE Deletion mutant delta 175-179 of human p55 TNF-R spacer region.
XX
KW p55; tumour necrosis factor receptor; TNF-R; human; murine; chimera;
KW epidermal growth factor receptor; EGF-R; protease; inhibitor;
KW phorbol myristate acetate; PMA; spacer region.
XX
OS Synthetic.
XX
PN AU9475742-A.
XX
PD 04-MAY-1995.
XX
PF 11-OCT-1994; 94AU-0075742.
XX
PR 12-OCT-1993; 93IL-0107268.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Batkin M, Brakebusch C, Varfolomeev E, Wallach D;
XX
DR WPI; 1995-194342/26.
XX
PT New protease capable of cleaving soluble tumour necrosis factor
PT (TNF) receptor - from cell-bound TNF- receptor, useful for
PT antagonising deleterious effects of TNF.
XX
PS Example 2; Fig 5; 40pp; English.
XX
CC The sequences represented in AAR75013-25 are fragments of deletion
CC mutants of the spacer region of human p55 tumour necrosis factor (TNF-R),
CC shown in AAR75012. Expression of this receptor is regulated by shedding
CC of the extracellular receptor fragment. The p55 TNF-R can be shed in
CC response to different inducing agents, e.g. phorbol myristate acetate
CC (PMA), depending on cell type. The only region of the receptor whose
CC structure affects the shedding response is the spacer region in the
CC extracellular domain. This region is located close to a site of cleavage
CC of the molecule, and links the Cys rich module to the transmembranal
CC domain. The sequences shown in AAR75026-47 are fragments of replacement
CC mutations of the spacer region. These mutations were introduced in order
CC to create an inhibitor of a protease that is capable of cleaving the
CC soluble TNF-R from the cell bound TNF-R. The relevant fragments of the
CC successful inhibitors can be seen in AAR75017-9, AAR75025, AAR75033-5 and
CC AAR75042-3. The protease inhibitors can be used for enhancing TNF
CC function.
XX
SQ Sequence 9 AA;

Query Match 39.3%; Score 22; DB 16; Length 9;
Best Local Similarity 55.6%; Pred. No. 6.4e+05;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
Qy      1 IDNVKKARV 9
          |:|||   |
Db      1 ienvkgttv 9
```

Search completed: July 8, 2002, 11:42:58
Job time: 230 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 8, 2002, 11:47:54 ; Search time 14.12 Seconds
(without alignments)
217.766 Million cell updates/sec

Title: US-09-461-061A-4
Perfect score: 161
Sequence: 1 TLTHTITKLNAENNATFYFKIDNVKKARVQVV 32

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 6873

Minimum DB seq length: 0
Maximum DB seq length: 32

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	34	21.1	31	2	G70123	hypothetical prote
2	33.5	20.8	20	2	D41299	T-cell receptor al
3	33	20.5	25	2	S03500	T-cell receptor ga
4	31	19.3	30	2	S77821	hypothetical prote
5	30.5	18.9	32	2	S32971	dodecandrin - endo
6	30	18.6	25	2	H64710	hypothetical prote
7	29	18.0	20	2	A53875	creatine kinase (E
8	29	18.0	26	2	S16336	beta-conglycinin b
9	28.5	17.7	29	2	I50214	protein-tyrosine-p
10	28	17.4	16	2	PH1580	Ig H chain V-D-J r
11	28	17.4	30	2	JS0645	diuretic hormone I
12	28	17.4	30	2	F89864	hypothetical prote
13	28	17.4	32	2	A95108	hypothetical prote
14	27.5	17.1	29	2	PQ0782	NADH dehydrogenase
15	27	16.8	14	2	S62374	alpha-1-antichymot
16	27	16.8	18	2	S58855	botulinum neurotox
17	27	16.8	24	2	C85846	unknown protein en
18	27	16.8	25	2	B47689	flagellar core pro
19	27	16.8	28	2	I40034	trpE protein - Bac
20	27	16.8	28	2	PL0005	pepsin A (EC 3.4.2
21	27	16.8	31	2	E87331	hypothetical prote
22	27	16.8	32	2	C49480	major immunophilin
23	27	16.8	32	2	A44906	L1 protein - human
24	26.5	16.5	31	2	E95151	hypothetical prote
25	26.5	16.5	32	2	F81385	hypothetical prote
26	26	16.1	16	2	I52226	aldehyde dehydroge
27	26	16.1	18	2	A42576	steroid receptor c
28	26	16.1	24	2	I61491	seed protein ws-9
29	26	16.1	25	2	S39391	calpain II heavy c
30	26	16.1	28	2	S77854	glutamate--tRNA li
31	26	16.1	28	2	S11618	ribosomal protein
32	26	16.1	28	2	S49924	stp protein (Baker
33	26	16.1	29	2	A45474	thrombospondin 2 -
34	26	16.1	31	2	T44925	hypothetical prote
35	26	16.1	31	2	A64630	hypothetical prote
36	25	15.5	10	2	PH0807	T-cell receptor al
37	25	15.5	15	2	A56970	GLYMA1 - soybean (
38	25	15.5	15	2	PH1610	Ig H chain V-D-J r
39	25	15.5	21	2	A60225	pyruvate dehydroge
40	25	15.5	22	2	S12391	fanE protein - Esc
41	25	15.5	24	2	C47689	flagellar core pro
42	25	15.5	28	1	G9BPSV	gene 9 protein - s
43	25	15.5	30	2	S27306	surface layer prot
44	25	15.5	31	2	A24730	Ig kappa chain V r
45	25	15.5	31	2	S78738	protein YOL038c-a

ALIGNMENTS

RESULT 1
G70123
hypothetical protein BB0191 - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C;Accession: G70123
R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A;Authors: Smith, H.O.; Venter, J.C.
A;Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A;Reference number: A70100; MUID:98065943
A;Accession: G70123
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-31 <KLE>
A;Cross-references: GB:AE001129; GB:AE000783; NID:g2688071; PIDN:AAC66583.1; PID:g268
A;Experimental source: strain B31

Query Match 21.1%; Score 34; DB 2; Length 31;
Best Local Similarity 41.7%; Pred. No. 4.6e+02;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

```
Qy      20 KIDNVKKARVQV 31
           ||:|: |  :|:
Db       7 KINNIDKTNIQI 18
```

RESULT 2
D41299
T-cell receptor alpha chain precursor (26.1) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 28-May-1992 #sequence_revision 28-May-1992 #text_change 05-Nov-1999
C;Accession: D41299
R;Uematsu, Y.; Wege, H.; Straus, A.; Ott, M.; Bannwarth, W.; Lanchbury, J.; Panayi, G
Proc. Natl. Acad. Sci. U.S.A. 88, 8534-8538, 1991
A;Title: The T-cell-receptor repertoire in the synovial fluid of a patient with rheum
A;Reference number: A41299; MUID:92020887
A;Accession: D41299
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-20 <UEM>
A;Cross-references: GB:S57442; NID:g236324; PIDN:AAB19959.1; PID:g236325
C;Keywords: T-cell receptor

Query Match 20.8%; Score 33.5; DB 2; Length 20;
Best Local Similarity 47.4%; Pred. No. 3.5e+02;
Matches 9; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

```
Qy      1 TLTHTITKLNAENNATFYF 19
          :||| || ||   ::  ||
Db      1 SLTH-ITALNKVVDSAVYF 18


RESULT   3
S03500
T-cell receptor gamma chain V region (FT6) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 30-May-1997
C;Accession: S03500
R;Heilig, J.S.; Tonegawa, S.
Nature 322, 836-840, 1986
A;Title: Diversity of murine gamma genes and expression in fetal and adult T lymphocytes
A;Reference number: S03498; MUID:86311315
A;Accession: S03500
A;Molecule type: mRNA
A;Residues: 1-25 <HEI>
A;Cross-references: EMBL:X04396
A;Note: the authors translated the codon AAG for residue 1 as Leu and GCC for residue 24
A;Note: this sequence was determined from the nonfunctional differentiated gene
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: T-cell receptor


  Query Match             20.5%;  Score 33;  DB 2;  Length 25;
  Best Local Similarity   46.2%;  Pred. No. 5e+02;
  Matches    6;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy     12 ENNATFYFKIDNV 24
          ::: |||  |:||
Db      1 KDDGTFYLIINNV 13


RESULT   4
S77821
hypothetical protein MC318 - Mycoplasma capricolum (fragment)
C;Species: Mycoplasma capricolum
C;Date: 09-Oct-1997 #sequence_revision 14-Nov-1997 #text_change 07-Dec-1999
C;Accession: S77821; S46915
R;Bork, P.; Ouzounis, C.; Casari, G.; Schneider, R.; Sander, C.; Dolan, M.; Gilbert, W.;
Mol. Microbiol. 16, 955-967, 1995
A;Title: Exploring the Mycoplasma capricolum genome: a minimal cell reveals its physiolo
A;Reference number: S77739; MUID:96059641
A;Accession: S77821
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-30 <BOR>
A;Cross-references: EMBL:Z33233
A;Experimental source: ATCC 27343
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
C;Genetics:
A;Genetic code: SGC3


  Query Match             19.3%;  Score 31;  DB 2;  Length 30;
  Best Local Similarity   26.3%;  Pred. No. 1.1e+03;
  Matches    5;  Conservative    6;  Mismatches    8;  Indels    0;  Gaps    0;

Qy     10 NAENNATFYFKIDNVKKAR 28
          |:   : ||  |:|: : :
Db      4 NSRQXSRFYHVIENLXEXK 22


RESULT   5
S32971
dodecandrin - endod pokeweed (fragment)
C;Species: Phytolacca dodecandra (endod pokeweed)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 07-May-1999
C;Accession: S32971
R;Ready, M.P.; Adams, R.P.; Robertus, J.D.
Biochim. Biophys. Acta 791, 314-319, 1984
A;Title: Dodecandrin, a new ribosome-inhibiting protein from Phytolacca dodecandra.
A;Reference number: S32971; MUID:85097746
A;Accession: S32971
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-32 <REA>
C;Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology


  Query Match             18.9%;  Score 30.5;  DB 2;  Length 32;
  Best Local Similarity   28.0%;  Pred. No. 1.4e+03;
  Matches    7;  Conservative    6;  Mismatches    9;  Indels    3;  Gaps    1;

Qy      1 TLTHTITKLNAENNATFYFKIDNVK 25
          |: : :      | |||   :||::
Db      3 TIIYNVGSTTISNYATF---MDNLR 24


RESULT   6
H64710
hypothetical protein HP1528 - Helicobacter pylori (strain 26695)
C;Species: Helicobacter pylori
C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999
C;Accession: H64710
R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R
 Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKe
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey,
Nature 388, 539-547, 1997
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser,
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A;Reference number: A64520; MUID:97394467
A;Accession: H64710
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-25 <TOM>
A;Cross-references: GB:AE000651; GB:AE000511; NID:g2314708; PIDN:AAD08577.1; PID:g231


  Query Match             18.6%;  Score 30;  DB 2;  Length 25;
  Best Local Similarity   44.4%;  Pred. No. 1.3e+03;
  Matches    8;  Conservative    6;  Mismatches    2;  Indels    2;  Gaps    2;

Qy     17 FYFK-ID-NVKKARVQVV 32
          |||: :| ::||  |:|:
Db      5 FYFRFLDYSLKKGLVKVI 22


RESULT   7
A53875
creatine kinase (EC 2.7.3.2) CK-MM - coho salmon (fragment)
C;Species: Oncorhynchus kisutch (coho salmon)
C;Date: 26-Sep-1994 #sequence_revision 18-Nov-1994 #text_change 11-Apr-1997
C;Accession: A53875
R;White, K.C.; Babbitt, P.C.; Buechter, D.D.; Kenyon, G.L.
J. Protein Chem. 11, 489-494, 1992
A;Title: The principal islet of the Coho salmon (Oncorhyncus kisutch) contains the BB
A;Reference number: A53875; MUID:93080727
A;Accession: A53875
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-20 <WHI>
A;Experimental source: Brockmann body, principal islet
A;Note: sequence extracted from NCBI backbone (NCBIP:120599)
C;Superfamily: creatine kinase; creatine kinase repeat homology
C;Keywords: phosphotransferase


  Query Match             18.0%;  Score 29;  DB 2;  Length 20;
  Best Local Similarity   38.5%;  Pred. No. 1.4e+03;
  Matches    5;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy     10 NAENNATFYFKID 22
          |  ||    ||::
Db      4 NTHNNFKLNFKVE 16
```

```
RESULT   8
S16336
beta-conglycinin beta chain - soybean (fragment)
C;Species: Glycine max (soybean)
C;Date: 21-Nov-1993 #sequence_revision 27-Feb-1997 #text_change 09-May-1997
C;Accession: S16336
R;Coates, J.B.; Medeiros, J.S.; Thanh, V.H.; Nielsen, N.C.
Arch. Biochem. Biophys. 243, 184-194, 1985
A;Title: Characterization of the subunits of beta-conglycinin.
A;Reference number: S16334; MUID:86049421
A;Accession: S16336
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-26 <COA>
C;Superfamily: glycinin


  Query Match             18.0%;  Score 29;  DB 2;  Length 26;
  Best Local Similarity   31.2%;  Pred. No. 1.8e+03;
  Matches    5;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       8 KLNAENNATFYFKIDN 23
            |:  : |  || :  |
Db       2 KVREDENNPFYLRSSN 17


RESULT   9
I50214
protein-tyrosine-phosphatase (EC 3.1.3.48) - chicken (fragment)
C;Species: Gallus gallus (chicken)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 23-Jul-1999
C;Accession: I50214
R;Stoker, A.W.
Mech. Dev. 46, 201-217, 1994
A;Title: Isoforms of a novel cell adhesion molecule-like protein tyrosine phosphatase ar
A;Reference number: I50212; MUID:95001563
A;Accession: I50214
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-29 <STO>
A;Cross-references: GB:L32782; NID:g485750; PIDN:AAA64462.1; PID:g485751
C;Genetics:
A;Gene: CRYPalpha3
C;Superfamily: leukocyte antigen-related protein; fibronectin type III repeat homology;
ogy
C;Keywords: phosphoric monoester hydrolase; tyrosine-specific phosphatase


  Query Match             17.7%;  Score 28.5;  DB 2;  Length 29;
  Best Local Similarity   31.8%;  Pred. No. 2.3e+03;
  Matches    7;  Conservative    6;  Mismatches    8;  Indels    1;  Gaps    1;

Qy      12 ENNATFYFKIDNVKK-ARVQVV 32
           |||:|    :  :::  || |:
Db       8 ENNSTIVVMLTKLREMGRVSVL 29


RESULT  10
PH1580
Ig H chain V-D-J region (wild-type clone 3) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999
C;Accession: PH1580
R;Levinson, D.A.; Campos-Torres, J.; Leder, P.
J. Exp. Med. 178, 317-329, 1993
A;Title: Molecular characterization of transgene-induced immunodeficiency in B-less mice
A;Reference number: PH1580; MUID:93301609
A;Accession: PH1580
A;Molecule type: DNA
A;Residues: 1-16 <LEV>
A;Experimental source: bone marrow pre-B lymphocyte
C;Keywords: immunoglobulin


  Query Match             17.4%;  Score 28;  DB 2;  Length 16;
  Best Local Similarity   57.1%;  Pred. No. 1.5e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      17 FYFKIDN 23
           ||: :||
Db      10 FYYAMDN 16


RESULT  11
JS0645
diuretic hormone II - tobacco hornworm
N;Alternate names: diuretic peptide II
C;Species: Manduca sexta (tobacco hornworm)
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 20-Apr-2001
C;Accession: JS0645
R;Blackburn, M.B.; Kingan, T.G.; Bodnar, W.; Shabanowitz, J.; Hunt, D.F.; Kempe, T.;
Biochem. Biophys. Res. Commun. 181, 927-932, 1991
A;Title: Isolation and identification of a new diuretic peptide from the tobacco horn
A;Reference number: JS0645; MUID:92109782
A;Accession: JS0645
A;Molecule type: protein
A;Residues: 1-30 <BLA>
C;Comment: This hormone is a factor which increases the rate of fluid excretion in th
C;Keywords: amidated carboxyl end
F;30/Modified site: amidated carboxyl end (Val) #status experimental


  Query Match             17.4%;  Score 28;  DB 2;  Length 30;
  Best Local Similarity   30.0%;  Pred. No. 2.8e+03;
  Matches    6;  Conservative    4;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2 LTHTITKLNAENNATFYFKI 21
           | |   :  |:||  |  ::
Db      11 LQHRYMEKVAQNNRNFLNRV 30


RESULT  12
F89864
hypothetical protein [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C;Accession: F89864
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of meticillin-resistant Stapylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: F89864
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-30 <KUR>
A;Cross-references: GB:BA000018; PID:g13700777; PIDN:BAB42073.1; GSPDB:GN00149
A;Experimental source: strain N315
C;Genetics:
A;Gene: SAS024


  Query Match             17.4%;  Score 28;  DB 2;  Length 30;
  Best Local Similarity   46.2%;  Pred. No. 2.8e+03;
  Matches    6;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 THTITKLNAENNA 15
           |  |:| |:  ||
Db      15 TQRISKRNSTGNA 27


RESULT  13
A95108
```

```
hypothetical protein SP0934 [imported] - Streptococcus pneumoniae (strain TIGR4)
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C;Accession: A95108
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: A95108
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-32 <KUR>
A;Cross-references: GB:AE005672; PIDN:AAK75058.1; PID:g14972409; GSPDB:GN00164; TIGR:SP4
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP0934

  Query Match             17.4%;  Score 28;  DB 2;  Length 32;
  Best Local Similarity   46.2%;  Pred. No. 3e+03;
  Matches    6;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      14 NATFYFKIDNVKK 26
           |  |: :|: |||
Db      19 NPYFFDRIEVVKK 31


RESULT  14
PQ0782
NADH dehydrogenase (EC 1.6.99.3) 31K chain - fava bean mitochondrion (fragment)
N;Alternate names: complex I 31K chain; NADH--ubiquinone reductase 31K chain
C;Species: mitochondrion Vicia faba (fava bean)
C;Date: 03-May-1994 #sequence_revision 07-Oct-1994 #text_change 17-Mar-1999
C;Accession: PQ0782
R;Leterme, S.; Boutry, M.
Plant Physiol. 102, 435-443, 1993
A;Title: Purification and preliminary characterization of mitochondrial complex I (NADH:
A;Reference number: PQ0775; MUID:94151437
A;Accession: PQ0782
A;Molecule type: protein
A;Residues: 1-29 <LET>
C;Comment: Complex I, mitochondrial NADH-ubiquiquinone reductase, is the first of the th
 ranging from 5K to 75K.
C;Comment: This enzyme catalyzes electron transfer from endogenous NADH to ubiquinone by
C;Genetics:
A;Genome: mitochondrion
C;Keywords: electron transfer; mitochondrion; oxidoreductase

  Query Match             17.1%;  Score 27.5;  DB 2;  Length 29;
  Best Local Similarity   27.6%;  Pred. No. 3.2e+03;
  Matches    8;  Conservative    3;  Mismatches   13;  Indels    5;  Gaps    1;

Qy       6 ITKLN-----AENNATFYFKIDNVKKARV 29
           :| ||      :||     :  |  |  |
Db       1 VTALNYHLDSPDNNPNLLHEFHNAXKKEV 29


RESULT  15
S62374
alpha-1-antichymotrypsin - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 24-Aug-1996 #sequence_revision 13-Mar-1997 #text_change 31-Mar-1997
C;Accession: S62374
R;Tsuda, M.; Sei, Y.; Ohkubo, T.; Yamamura, M.; Kamiguchi, H.; Akatsuka, A.; Tsuda, T.;
Eur. J. Biochem. 235, 821-827, 1996
A;Title: The defective secretion of a naturally occurring alpha-1-antichymotrypsin varia
A;Reference number: S62374; MUID:96184564
A;Accession: S62374
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-14 <TSU>

  Query Match             16.8%;  Score 27;  DB 2;  Length 14;
  Best Local Similarity   42.9%;  Pred. No. 1.8e+03;
  Matches    6;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      14 NATFYFKIDNVKKA 27
           |  |  |: | |:|
Db       1 NIFFMSKVTNPKQA 14


RESULT  16
S58855
botulinum neurotoxin type B nontoxic-nonhemagglutinin component - Clostridium botulin
N;Alternate names: NTNH protein
C;Species: Clostridium botulinum
A;Variety: strain Eklund 17B
C;Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 15-Oct-1999
C;Accession: S58855
R;East, A.K.; Stacey, J.M.; Collins, M.D.
Syst. Appl. Microbiol. 17, 306-312, 1994
A;Title: Cloning and sequencing of a hemagglutinin component of the botulinum neuroto
A;Reference number: S58855
A;Accession: S58855
A;Molecule type: DNA
A;Residues: 1-18 <EAS>
A;Cross-references: EMBL:X79103; NID:g870932; PIDN:CAA55712.1; PID:g870933
A;Experimental source: strain Eklund 17B
C;Keywords: neurotoxin

  Query Match             16.8%;  Score 27;  DB 2;  Length 18;
  Best Local Similarity   29.4%;  Pred. No. 2.3e+03;
  Matches    5;  Conservative    4;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       9 LNAENNATFYFKIDNVK 25
           :|  :| :    :|| |
Db       1 MNINDNLSINSPVDNKK 17


RESULT  17
C85846
unknown protein encoded within prophage CP-933V [imported] - Escherichia coli (strain
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: C85846
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: C85846
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-24 <STO>
A;Cross-references: GB:AE005174; NID:g12516398; PIDN:AAG57231.1; GSPDB:GN00145; UWGP:
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z3347

  Query Match             16.8%;  Score 27;  DB 2;  Length 24;
  Best Local Similarity   25.0%;  Pred. No. 3.1e+03;
  Matches    5;  Conservative    7;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       4 HTITKLNAENNATFYFKIDN 23
           |   :| :: |:: | : |:
Db       4 HCCAELMSDPNSSMYEEEDD 23


RESULT  18
```

B47689
flagellar core protein, 37K - Treponema hyodysenteriae (fragment)
C;Species: Treponema hyodysenteriae
C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 24-Feb-1995
C;Accession: B47689
R;Koopman, M.B.; Baats, E.; van Vorstenbosch, C.J.; van der Zeijst, B.A.; Kusters, J.G.
J. Gen. Microbiol. 138, 2697-2706, 1992
A;Title: The periplasmic flagella of Serpulina (Treponema) hyodysenteriae are composed o
A;Reference number: A47689; MUID:93139764
A;Contents: C5, Treponema
A;Accession: B47689
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-25 <KOO>
A;Note: sequence extracted from NCBI backbone (NCBIP:123401)

Query Match 16.8%; Score 27; DB 2; Length 25;
Best Local Similarity 28.0%; Pred. No. 3.2e+03;
Matches 7; Conservative 8; Mismatches 8; Indels 2; Gaps 1;

```
Qy      2 LTHTITKLNAENNATFYFKIDNVKK 26
          : : |: :||:   |  |:  ::||
Db      3 INNNISAINAQR--TLKFRQVDLKK 25
```

RESULT 19
I40034
trpE protein - Bacillus pumilus (fragment)
C;Species: Bacillus pumilus
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 22-Jun-1999
C;Accession: I40034
R;Kuroda, M.I.; Shimotsu, H.; Henner, D.J.; Yanofsky, C.
J. Bacteriol. 167, 792-798, 1986
A;Title: Regulatory elements common to the Bacillus pumilus and Bacillus subtilis trp op
A;Reference number: I40034; MUID:86304167
A;Accession: I40034
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-28 <RES>
A;Cross-references: GB:M14027; NID:g143781; PIDN:AAA22872.1; PID:g551734
C;Superfamily: anthranilate synthase component I

Query Match 16.8%; Score 27; DB 2; Length 28;
Best Local Similarity 33.3%; Pred. No. 3.6e+03;
Matches 4; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

```
Qy      9 LNAENNATFYFK 20
          :|:::| | : |
Db      1 MNSQSNLTQFLK 12
```

RESULT 20
PL0005
pepsin A (EC 3.4.23.1) precursor - Mongolian sheep (fragments)
C;Species: Ovis platyurea (Mongolian sheep)
C;Date: 30-Jun-1992 #sequence_revision 19-Jan-2001 #text_change 19-Jan-2001
C;Accession: PL0005
R;Baudys, M.; Erdene, T.G.; Kostka, V.; Pavlik, M.; Foltmann, B.
Comp. Biochem. Physiol. B 89, 385-391, 1988
A;Title: Comparison between prochymosin and pepsinogen from lamb and calf.
A;Reference number: PL0006; MUID:88185059
A;Accession: PL0005
A;Molecule type: protein
A;Residues: 1-15;16-28 <BAU>
A;Experimental source: stomach
C;Keywords: aspartic proteinase; hydrolase; protein digestion; zymogen
F;1-15/Domain: activation peptide (fragment) #status predicted <ACP>

Query Match 16.8%; Score 27; DB 2; Length 28;
Best Local Similarity 50.0%; Pred. No. 3.6e+03;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```
Qy     19 FKIDNVKKARVQVV 32
          |||  |||  :: |
Db      3 FKIPLVKKKSLRQV 16
```

Search completed: July 8, 2002, 11:50:14
Job time: 140 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 8, 2002, 11:49:54 ; Search time 10.26 Seconds
(without alignments)
120.763 Million cell updates/sec

Title: US-09-461-061A-4
Perfect score: 161
Sequence: 1 TLTHTITKLNAENNATFYFKIDNVKKARVQVV 32

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 2064

Minimum DB seq length: 0
Maximum DB seq length: 32

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	34	21.1	24	1	DNAJ_STRAG	P95694	streptococc
2	34	21.1	31	1	Y191_BORBU	O51209	borrelia bu
3	28.5	17.7	20	1	CHP_THICU	P80486	thiobacillu
4	28	17.4	30	1	DIU2_HYLLI	P82015	hyles linea
5	28	17.4	30	1	DIU2_MANSE	P24858	manduca sex
6	27	16.8	25	1	FLB1_TREHY	P80158	treponema h
7	25	15.5	15	1	MM01_RAT	P81563	rattus norv
8	25	15.5	19	1	TRPB_KLEAE	P14552	klebsiella
9	25	15.5	20	1	FLA2_SPIAU	P21985	spirochaeta
10	25	15.5	27	1	FLA1_SPIAU	P21984	spirochaeta
11	25	15.5	28	1	VG9_SPV4	P11341	spiroplasma
12	25	15.5	31	1	COX4_NEUCR	P06809	neurospora
13	24.5	15.2	18	1	FMF1_ECOLI	P20860	escherichia
14	24.5	15.2	29	1	RS7_METTE	O93639	methanosarc
15	24	14.9	15	1	NIS1_ANAVA	Q44507	anabaena va
16	24	14.9	21	1	FA71_TETPY	P81424	tetrahymena
17	24	14.9	21	1	LPT_ECOLI	P03059	escherichia
18	24	14.9	24	1	LPTN_ECOLI	P09408	escherichia
19	24	14.9	26	1	PUTA_KLEPN	P23725	klebsiella
20	24	14.9	30	1	AMPT_BACST	P00728	bacillus st
21	24	14.9	31	1	CEC1_PIG	P14661	sus scrofa
22	23.5	14.6	29	1	AL21_HORSE	P81216	equus cabal
23	23.5	14.6	30	1	RIPS_MOMCO	P20655	momordica c
24	23.5	14.6	32	1	DBH_SYNY1	P02343	synechocyst
25	23	14.3	19	1	FLA3_SPIAU	P21986	spirochaeta
26	23	14.3	20	1	LPP2_HUMAN	P56642	homo sapien
27	23	14.3	21	1	CXGT_CONTU	P17684	conus tulip
28	23	14.3	27	1	FLA4_SPIAU	P21987	spirochaeta
29	23	14.3	31	1	HEM2_PHAGO	P27687	phascolopsi
30	23	14.3	32	1	YCPG_MASLA	P29735	mastigoclad
31	22	13.7	16	1	ALL1_CALVO	P41839	calliphora
32	22	13.7	24	1	HEMU_LINRE	P23544	lingula ree
33	22	13.7	26	1	YFHA_KLEPN	P21710	klebsiella
34	22	13.7	28	1	ETX2_BACCE	P80568	bacillus ce
35	22	13.7	29	1	MDH_BURPS	P80536	burkholderi
36	22	13.7	30	1	IHFB_RHILE	P80606	rhizobium l
37	22	13.7	32	1	RIP2_PHYDI	P34967	phytolacca
38	21	13.0	13	1	LIGA_TRAVE	P20011	trametes ve
39	21	13.0	13	1	RPOC_MYCGA	P47716	mycoplasma
40	21	13.0	17	1	B29K_PORGI	P81784	porphyromon
41	21	13.0	19	1	TX3_PHONI	P31010	phoneutria
42	21	13.0	20	1	DCMS_PSECF	P19915	pseudomonas
43	21	13.0	20	1	MCRG_METTE	P22950	methanosarc
44	21	13.0	21	1	CSPS_STRTR	P81622	streptococc
45	21	13.0	21	1	ECD1_LYMDI	P80936	lymantria d

ALIGNMENTS

RESULT 1
DNAJ_STRAG
ID DNAJ_STRAG STANDARD; PRT; 24 AA.
AC P95694;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chaperone protein dnaJ (Fragment).
GN DNAJ.
OS Streptococcus agalactiae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1311;
RN [1]
RP SEQUENCE FROM N.A.
RA Rioux C.R., Martin D., Hamel J., Brodeur B.R.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: ACTS AS A CO-CHAPERONE. STIMULATES, JOINTLY WITH GRPE,
CC THE ATPASE ACTIVITY OF DNAK (BY SIMILARITY).
CC -!- COFACTOR: BINDS TWO ZINC IONS PER MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE DNAJ FAMILY.
CC -!- SIMILARITY: CONTAINS 1 J DOMAIN.
CC ---

CC ---
DR EMBL; U72719; AAB39220.1; -.
DR InterPro; IPR001305; DnaJ_CXXCXGXG.
DR InterPro; IPR001623; DnaJ_N.
DR PROSITE; PS00636; DNAJ_1; PARTIAL.
DR PROSITE; PS50076; DNAJ_2; PARTIAL.
DR PROSITE; PS00637; DNAJ_CXXCXGXG; PARTIAL.
KW Chaperone; DNA replication; Heat shock; Repeat; Zinc; Metal-binding.
FT DOMAIN 3 >24 J-DOMAIN.
FT NON_TER 24 24
SQ SEQUENCE 24 AA; 2760 MW; 342AE656E00913FC CRC64;

Query Match 21.1%; Score 34; DB 1; Length 24;
Best Local Similarity 34.8%; Pred. No. 83;
Matches 8; Conservative 3; Mismatches 4; Indels 8; Gaps 1;

```
Qy      13 NNATFYFKI--------DNVKKA 27
           ||  || ::         | :|||
Db       2 NNTEFYDRLGVSKDASQDEIKKA 24
```

RESULT 2
Y191_BORBU
ID Y191_BORBU STANDARD; PRT; 31 AA.

```
AC   O51209;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein BB0191.
GN   BB0191.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=98065943; PubMed=9403685;
RA   Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA   Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA   Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA   Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA   van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA   Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA   Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA   Smith H.O., Venter J.C.;
RT   "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT   burgdorferi.";
RL   Nature 390:580-586(1997).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001129; AAC66583.1; -.
DR   TIGR; BB0191; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   31 AA;  3583 MW;  4DDA4AF26162455E CRC64;


  Query Match             21.1%;  Score 34;  DB 1;  Length 31;
  Best Local Similarity   41.7%;  Pred. No. 1.1e+02;
  Matches    5;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy     20 KIDNVKKARVQV 31
          ||:|: |  :|:
Db      7 KINNIDKTNIQI 18


RESULT   3
CHP_THICU
ID   CHP_THICU      STANDARD;      PRT;    20 AA.
AC   P80486;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   Chemoheterotroph-specific protein (Fragment).
OS   Thiobacillus cuprinus.
OC   Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Thiomonas.
OX   NCBI_TaxID=36860;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=DSM 5494;
RA   Marin I., Amaro A.M., Jerez C.A., Amils R., Abad J.P.;
RL   Submitted (SEP-1995) to the SWISS-PROT data bank.
CC   -!- MISCELLANEOUS: FOUND SPECIFICALLY IN CELLS CULTURED
CC       CHEMOHETEROTROPHICALLY.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2167 MW;  E5BAEC1BA3238A0A CRC64;


  Query Match             17.7%;  Score 28.5;  DB 1;  Length 20;
  Best Local Similarity   42.9%;  Pred. No. 4e+02;
  Matches    9;  Conservative    3;  Mismatches    6;  Indels    3;  Gaps    1;

Qy      2 LTHTITKLNAENNATFYFKID 22
          ||:| | :| :   ||  | |
Db      2 LTYTGTVINVQ---TFAAKPD 19


RESULT   4
DIU2_HYLLI
ID   DIU2_HYLLI     STANDARD;      PRT;    30 AA.
AC   P82015;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Diuretic hormone 2 (DH-2) (Diuretic peptide 2) (DP-2) (DH(30)).
OS   Hyles lineata (Whitelined sphinx moth).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Sphingiodea; Sphingidae; Macroglossinae; Hyles.
OX   NCBI_TaxID=103890;
RN   [1]
RP   SEQUENCE, ACTIVITY, AND AMIDATION.
RC   TISSUE=Head;
RX   MEDLINE=20161578; PubMed=10696588;
RA   Furuya K., Harper M.A., Schegg K.M., Shooley D.A.;
RT   "Isolation and characterization of CFR-related diuretic hormones from
RT   the whitelined sphinx moth Hyles lineata.";
RL   Insect Biochem. Mol. Biol. 20:127-133(2000).
CC   -!- FUNCTION: REGULATION OF FLUID SECRETION. MAY STIMULATE PRIMARY
CC       URINE SECRETION BY MALPIGHIAN TUBULES AND CAUSES A DOSE-DEPENDENT
CC       STIMULATION OF CAMP LEVELS IN THE TUBULES.
CC   -!- MASS SPECTROMETRY: MW=3574; METHOD=MALDI.
CC   -!- SIMILARITY: BELONGS TO THE SAUVAGINE/CORTICOTROPIN-RELEASING
CC       FACTOR/UROTENSIN I FAMILY OF PEPTIDES.
DR   InterPro; IPR000187; CRF.
DR   Pfam; PF00473; CRF; 1.
DR   SMART; SM00039; CRF; 1.
DR   PROSITE; PS00511; CRF; FALSE_NEG.
KW   Hormone; Amidation.
FT   MOD_RES      30     30       AMIDATION.
SQ   SEQUENCE   30 AA;  3575 MW;  5C2D6BD2DD8BFC67 CRC64;


  Query Match             17.4%;  Score 28;  DB 1;  Length 30;
  Best Local Similarity   30.0%;  Pred. No. 7.4e+02;
  Matches    6;  Conservative    4;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      2 LTHTITKLNAENNATFYFKI 21
          | |   :  |:|| | |  ::
Db     11 LQHRYMEKVAQNNRNFLNRV 30


RESULT   5
DIU2_MANSE
ID   DIU2_MANSE     STANDARD;      PRT;    30 AA.
AC   P24858;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Diuretic hormone 2 (DH-2) (Diuretic peptide 2) (DP-2) (DPII).
OS   Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Sphingiodea; Sphingidae; Sphinginae; Manduca.
OX   NCBI_TaxID=7130;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Corpora cardiaca;
RX   MEDLINE=92109782; PubMed=1764106;
RA   Blackburn M.B., Kingan T.G., Bodnar W., Shabanowitz J., Hunt D.F.,
RA   Kempe T., Wagner R.M., Raina A.K., Schnee M.E., Ma M.C.;
RT   "Isolation and identification of a new diuretic peptide from the
RT   tobacco hornworm, Manduca sexta.";
RL   Biochem. Biophys. Res. Commun. 181:927-932(1991).
```

```
CC   -!- FUNCTION: REGULATION OF FLUID SECRETION.
CC   -!- SIMILARITY: BELONGS TO THE SAUVAGINE/CORTICOTROPIN-RELEASING
CC       FACTOR/UROTENSIN I FAMILY OF PEPTIDES.
DR   PIR; JS0645; JS0645.
DR   InterPro; IPR000187; CRF.
DR   Pfam; PF00473; CRF; 1.
DR   SMART; SM00039; CRF; 1.
DR   PROSITE; PS00511; CRF; FALSE_NEG.
KW   Hormone; Amidation.
FT   MOD_RES      30     30       AMIDATION.
SQ   SEQUENCE   30 AA;  3561 MW;  5C2D6BD2DD8A2717 CRC64;


  Query Match             17.4%;  Score 28;  DB 1;  Length 30;
  Best Local Similarity   30.0%;  Pred. No. 7.4e+02;
  Matches    6;  Conservative    4;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       2 LTHTITKLNAENNATFYFKI 21
           | |   :  |:||  |  ::
Db      11 LQHRYMEKVAQNNRNFLNRV 30


RESULT   6
FLB1_TREHY
ID   FLB1_TREHY     STANDARD;      PRT;    25 AA.
AC   P80158;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Flagellar filament core protein flaB1 (37 kDa core protein)
DE   (Fragment).
GN   FLAB1.
OS   Treponema hyodysenteriae (Serpulina hyodysenteriae).
OC   Bacteria; Spirochaetales; Brachyspiraceae; Brachyspira.
OX   NCBI_TaxID=159;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=C5;
RX   MEDLINE=93139764; PubMed=1487733;
RA   Koopman M.B.H., Baats E., van Vorstenbosch C.J.A.H.V.,
RA   van der Zeijst B.A.M., Kusters J.G.;
RT   "The periplasmic flagella of Serpulina (Treponema) hyodysenteriae are
RT   composed of two sheath proteins and three core proteins.";
RL   J. Gen. Microbiol. 138:2697-2706(1992).
CC   -!- FUNCTION: COMPONENT OF THE CORE OF THE FLAGELLA.
CC   -!- SUBUNIT: THE FLAGELLUM CONSISTS OF AN OUTER LAYER COMPOSED OF TWO
CC       SHEATH PROTEINS, FLAA1 (44 kDa) AND FLAA2 (35 kDa) AROUND A CORE
CC       THAT CONTAINS THREE PROTEINS FLAB1 (37 kDa), FLAB2 (34 kDa) AND
CC       FLAB3 (32 kDa).
CC   -!- SUBCELLULAR LOCATION: Periplasmic flagellum.
CC   -!- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.
DR   PIR; B47689; B47689.
KW   Flagella; Periplasmic.
FT   NON_TER      25     25
SQ   SEQUENCE   25 AA;  2915 MW;  BBB699593CD398B6 CRC64;


  Query Match             16.8%;  Score 27;  DB 1;  Length 25;
  Best Local Similarity   28.0%;  Pred. No. 8.3e+02;
  Matches    7;  Conservative    8;  Mismatches    8;  Indels    2;  Gaps    1;

Qy       2 LTHTITKLNAENNATFYFKIDNVKK 26
           : : |: :||:   |  |:  ::||
Db       3 INNNISAINAQR--TLKFRQVDLKK 25


RESULT   7
MM01_RAT
ID   MM01_RAT       STANDARD;      PRT;    15 AA.
AC   P81563;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Interstitial collagenase (EC 3.4.24.7) (Matrix metalloproteinase-1)
DE   (MMP-1) (Fibroblast collagenase) (Myocardial collagenase) (Fragment).
GN   MMP1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Heart;
RX   MEDLINE=96201136; PubMed=8605638;
RA   Tyagi S.C., Cleutjens J.P.M.;
RT   "Myocardial collagenase: purification and structural
RT   characterization.";
RL   Can. J. Cardiol. 12:165-171(1996).
CC   -!- FUNCTION: CLEAVES COLLAGENS OF TYPES I, II, AND III AT ONE SITE IN
CC       THE HELICAL DOMAIN. ALSO CLEAVES COLLAGENS OF TYPES VII AND X. MAY
CC       PLAY A ROLE IN THE DETERIORATION OF THE HEART WALL EXTRACELLULAR
CC       MATRIX PROTEINS DURING THE ONSET OF DIALATED CARDIOMYOPATHY.
CC   -!- CATALYTIC ACTIVITY: Cleaves preferentially one bond in native
CC       collagen. Cleavage of the triple helix of collagen at about three-
CC       quarters of the length of the molecule from the N-terminus, at
CC       775-Gly-|-Ile-776 in the alpha-1(I) chain. Cleaves synthetic
CC       substrates and alpha-macroglobulins at bonds where P1' is a
CC       hydrophobic residue.
CC   -!- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
CC   -!- ENZYME REGULATION: CAN BE ACTIVATED WITHOUT REMOVAL OF THE
CC       ACTIVATION PEPTIDE.
CC   -!- PTM: THE N-TERMINAL IS BLOCKED.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
CC       METALLOPROTEASE) ALSO KNOWN AS MATRIXIN SUBFAMILY.
KW   Hydrolase; Metalloprotease; Zinc; Calcium; Collagen degradation;
KW   Extracellular matrix.
FT   NON_TER       1      1
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1787 MW;  15A57D24C0F6FD80 CRC64;


  Query Match             15.5%;  Score 25;  DB 1;  Length 15;
  Best Local Similarity   55.6%;  Pred. No. 9.2e+02;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       9 LNAENNATF 17
           | :| || |
Db       3 LKSEKNADF 11


RESULT   8
TRPB_KLEAE
ID   TRPB_KLEAE     STANDARD;      PRT;    19 AA.
AC   P14552;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Tryptophan synthase beta chain (EC 4.2.1.20) (Fragment).
GN   TRPB.
OS   Klebsiella aerogenes.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Klebsiella.
OX   NCBI_TaxID=28451;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81199002; PubMed=6262736;
RA   Nichols B.P., Blumenberg M., Yanofsky C.;
RT   "Comparison of the nucleoside sequence of trpA and sequences
RT   immediately beyond the trp operon of Klebsiella aerogenes. Salmonella
RT   typhimurium and Escherichia coli.";
RL   Nucleic Acids Res. 9:1743-1755(1981).
CC   -!- FUNCTION: THE BETA SUBUNIT IS RESPONSIBLE FOR THE SYNTHESIS
CC       OF L-TRYPTOPHAN FROM INDOLE AND L-SERINE.
CC   -!- CATALYTIC ACTIVITY: L-serine + 1-(indol-3-yl)glycerol 3-phosphate
CC       = L-tryptophan + glyceraldehyde 3-phosphate + H(2)O.
CC   -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
```

```
CC   -!- PATHWAY: LAST (FIFTH) STEP IN BIOSYNTHESIS OF TRYPTOPHAN.
CC   -!- SUBUNIT: TETRAMER OF TWO ALPHA AND TWO BETA CHAINS (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE TRPB FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V00630; CAA23901.1; -.
DR   EMBL; J01738; AAA25144.1; -.
DR   HSSP; P00933; 2TYS.
DR   InterPro; IPR000993; Trp_synth_beta.
DR   PROSITE; PS00168; TRP_SYNTHASE_BETA; PARTIAL.
KW   Tryptophan biosynthesis; Pyridoxal phosphate; Lyase.
FT   NON_TER       1      1
SQ   SEQUENCE   19 AA;  2183 MW;  82864627BF574E2C CRC64;


  Query Match            15.5%;  Score 25;  DB 1;  Length 19;
  Best Local Similarity   30.8%;  Pred. No. 1.2e+03;
  Matches    4;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy     19 FKIDNVKKARVQV 31
          | : :: ||| ::
Db      7 FTVHDILKARGEI 19


RESULT   9
FLA2_SPIAU
ID   FLA2_SPIAU     STANDARD;      PRT;     20 AA.
AC   P21985;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Flagellar filament 33 kDa core protein (Minor) (Fragment).
OS   Spirochaeta aurantia.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Spirochaeta.
OX   NCBI_TaxID=147;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=M1;
RX   MEDLINE=91123217; PubMed=1991729;
RA   Parales J. Jr., Greenberg E.P.;
RT   "N-terminal amino acid sequences and amino acid compositions of the
RT   Spirochaeta aurantia flagellar filament polypeptides.";
RL   J. Bacteriol. 173:1357-1359(1991).
CC   -!- FUNCTION: COMPONENT OF THE CORE OF THE FLAGELLA.
CC   -!- SUBUNIT: THE FLAGELLUM CONSISTS OF AN OUTER LAYER COMPOSED OF
CC       REPEATING UNITS OF FLAA AROUND A CORE THAT CONTAINS ONE OR ALL OF
CC       FIVE ANTIGENICALLY RELATED POLYPEPTIDES.
CC   -!- SUBCELLULAR LOCATION: Periplasmic flagellum.
CC   -!- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.
KW   Flagella; Periplasmic.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2166 MW;  FFCE66276472BF5B CRC64;


  Query Match            15.5%;  Score 25;  DB 1;  Length 20;
  Best Local Similarity   27.3%;  Pred. No. 1.2e+03;
  Matches    3;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 LTHTITKLNAE 12
          : | :: :||:
Db      3 INHNMSAINAQ 13


RESULT  10
FLA1_SPIAU
ID   FLA1_SPIAU     STANDARD;      PRT;     27 AA.
AC   P21984;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Flagellar filament 34 kDa core protein (Major) (Fragment).
OS   Spirochaeta aurantia.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Spirochaeta.
OX   NCBI_TaxID=147;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=M1;
RX   MEDLINE=91123217; PubMed=1991729;
RA   Parales J. Jr., Greenberg E.P.;
RT   "N-terminal amino acid sequences and amino acid compositions of the
RT   Spirochaeta aurantia flagellar filament polypeptides.";
RL   J. Bacteriol. 173:1357-1359(1991).
CC   -!- FUNCTION: COMPONENT OF THE CORE OF THE FLAGELLA.
CC   -!- SUBUNIT: THE FLAGELLUM CONSISTS OF AN OUTER LAYER COMPOSED OF
CC       REPEATING UNITS OF FLAA AROUND A CORE THAT CONTAINS ONE OR ALL OF
CC       FIVE ANTIGENICALLY RELATED POLYPEPTIDES.
CC   -!- SUBCELLULAR LOCATION: Periplasmic flagellum.
CC   -!- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.
KW   Flagella; Periplasmic.
FT   NON_TER      27     27
SQ   SEQUENCE   27 AA;  2942 MW;  CEB9060FA32682BF CRC64;


  Query Match            15.5%;  Score 25;  DB 1;  Length 27;
  Best Local Similarity   27.3%;  Pred. No. 1.7e+03;
  Matches    3;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 LTHTITKLNAE 12
          : | :: :||:
Db      3 INHNMSAINAQ 13


RESULT  11
VG9_SPV4
ID   VG9_SPV4       STANDARD;      PRT;     28 AA.
AC   P11341;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-JUL-1989 (Rel. 11, Last sequence update)
DT   01-JUL-1989 (Rel. 11, Last annotation update)
DE   Gene 9 protein.
GN   9.
OS   Spiroplasma virus 4 (SpV4).
OC   Viruses; ssDNA viruses; Microviridae; Spiromicrovirus.
OX   NCBI_TaxID=10855;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88032809; PubMed=2822658;
RA   Renaudin J., Pascarel M.-C., Bove J.-M.;
RT   "Spiroplasma virus 4: nucleotide sequence of the viral DNA,
RT   regulatory signals, and proposed genome organization.";
RL   J. Bacteriol. 169:4950-4961(1987).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M17988; -; NOT_ANNOTATED_CDS.
DR   PIR; B29825; G9BPSV.
SQ   SEQUENCE   28 AA;  3776 MW;  9916C3C9C3B9FD1D CRC64;


  Query Match            15.5%;  Score 25;  DB 1;  Length 28;
  Best Local Similarity   50.0%;  Pred. No. 1.8e+03;
  Matches    6;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;
```

```
Qy      19 FKIDNVKKARVQ 30
             ||   | | ||:
Db       6 FKTSRVVKHRVR 17


RESULT  12
COX4_NEUCR
ID   COX4_NEUCR     STANDARD;      PRT;    31 AA.
AC   P06809;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Cytochrome c oxidase polypeptide IV, mitochondrial precursor
DE   (EC 1.9.3.1) (Fragment).
OS   Neurospora crassa.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Sordariales; Sordariaceae; Neurospora.
OX   NCBI_TaxID=5141;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86085927; PubMed=3001085;
RA   Sachs M.S., David M., Werner S., Rajbhandary U.L.;
RT   "Nuclear genes for cytochrome c oxidase subunits of Neurospora
RT   crassa. Isolation and characterization of cDNA clones for subunits
RT   IV, V, VI, and possibly VII.";
RL   J. Biol. Chem. 261:869-873(1986).
CC   -!- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC       c + 2 H(2)O.
CC   -!- SUBCELLULAR LOCATION: Mitochondrial inner membrane.
CC   -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VB FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M12116; AAA33574.1; -.
DR   PIR; A25629; A25629.
KW   Oxidoreductase; Mitochondrion; Inner membrane; Transit peptide.
FT   NON_TER       1      1
FT   TRANSIT      <1     16       MITOCHONDRION.
FT   CHAIN        17    >31       CYTOCHROME C OXIDASE POLYPEPTIDE IV.
FT   NON_TER      31     31
SQ   SEQUENCE   31 AA;  3397 MW;  069F5D5510300362 CRC64;


  Query Match             15.5%;  Score 25;  DB 1;  Length 31;
  Best Local Similarity   33.3%;  Pred. No. 2e+03;
  Matches    4;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKLNAE 12
           ::  |: : |||
Db       8 SIATTVVRCNAE 19


RESULT  13
FMF1_ECOLI
ID   FMF1_ECOLI     STANDARD;      PRT;    18 AA.
AC   P20860;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-FEB-1991 (Rel. 17, Last annotation update)
DE   F7-1 fimbrial protein (F7-1 pilin) (P adhesin) (Fragment).
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=89079313; PubMed=2562836;
RA   Hoschuetzky H., Lottspeich F., Jann K.;
RT   "Isolation and characterization of the alpha-galactosyl-1,4-beta-
RT   galactosyl-specific adhesin (P adhesin) from fimbriated Escherichia
RT   coli.";
RL   Infect. Immun. 57:76-81(1989).
CC   -!- FUNCTION: FIMBRIAE (ALSO CALLED PILI), POLAR FILAMENTS RADIATING
CC       FROM THE SURFACE OF THE BACTERIUM TO A LENGTH OF 0.5-1.5
CC       MICROMETERS AND NUMBERING 100-300 PER CELL, ENABLE BACTERIA TO
CC       COLONIZE THE EPITHELIUM OF SPECIFIC HOST ORGANS.
CC   -!- MISCELLANEOUS: THIS IS A ALPHA-GALACTOSYL-1,4-BETA-GALACTOSYL-SPE
CC       CIFIC ADHESIN.
DR   PIR; A30541; A30541.
KW   Fimbria.
FT   NON_TER       18     18
SQ   SEQUENCE   18 AA;  1956 MW;  E9A44CDA6E2886B5 CRC64;


  Query Match             15.2%;  Score 24.5;  DB 1;  Length 18;
  Best Local Similarity   50.0%;  Pred. No. 1.3e+03;
  Matches    6;  Conservative    1;  Mismatches    4;  Indels    1;  Gaps    1;

Qy      13 NNATFYFKIDNV 24
             ||  ||  : ||
Db       2 NNIVFY-SLGNV 12


RESULT  14
RS7_METTE
ID   RS7_METTE      STANDARD;      PRT;    29 AA.
AC   O93639;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   30S ribosomal protein S7P (Fragment).
GN   RPS7P OR S7.
OS   Methanosarcina thermophila.
OC   Archaea; Euryarchaeota; Methanosarcinales; Methanosarcinaceae;
OC   Methanosarcina.
OX   NCBI_TaxID=2210;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM 1825 / TM-1;
RX   MEDLINE=99059471; PubMed=9845338;
RA   Thomas T., Cavicchioli R.;
RT   "Archaeal cold-adapted proteins: structural and evolutionary analysis
RT   of the elongation factor 2 proteins from psychrophilic, mesophilic and
RT   thermophilic methanogens.";
RL   FEBS Lett. 439:281-287(1998).
CC   -!- FUNCTION: PROTEIN S7 BINDS SPECIFICALLY TO PART OF THE 3' END OF
CC       16S RIBOSOMAL RNA (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE S7P FAMILY OF RIBOSOMAL PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF026165; AAC79199.1; -.
DR   InterPro; IPR000235; Ribosomal_S7.
DR   PROSITE; PS00052; RIBOSOMAL_S7; PARTIAL.
KW   Ribosomal protein; rRNA-binding.
FT   NON_TER       1      1
SQ   SEQUENCE   29 AA;  3217 MW;  1602B8A2E6C50C2B CRC64;


  Query Match             15.2%;  Score 24.5;  DB 1;  Length 29;
  Best Local Similarity   39.1%;  Pred. No. 2.2e+03;
  Matches    9;  Conservative    3;  Mismatches   10;  Indels    1;  Gaps    1;
```

```
Qy       9 LNAENNATFYFKIDNVKKARVQV 31
           : | |  |  | | | | |: :|
Db       3 IGAANRDTKSFSI-NRKDAKERV 24


RESULT  15
NIS1_ANAVA
ID   NIS1_ANAVA     STANDARD;      PRT;    15 AA.
AC   Q44507;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Cysteine desulfurase 1 (EC 4.4.1.-) (Nitrogenase metalloclusters
DE   biosynthesis protein nifS1) (Fragment).
GN   NIFS1 OR NIFS.
OS   Anabaena variabilis.
OC   Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.
OX   NCBI_TaxID=1172;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PCC 7937 / ATCC 29413;
RA   Monnerjahn U., Boehme H.;
RL   Submitted (DEC-1992) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   EXPRESSION PATTERN.
RC   STRAIN=PCC 7937 / ATCC 29413;
RX   MEDLINE=96016168; PubMed=7568132;
RA   Thiel T., Lyons E.M., Erker J.C., Ernst A.;
RT   "A second nitrogenase in vegetative cells of a heterocyst-forming
RT   cyanobacterium.";
RL   Proc. Natl. Acad. Sci. U.S.A. 92:9358-9362(1995).
CC   -!- FUNCTION: CATALYZES THE REMOVAL OF ELEMENTAL SULFUR ATOMS FROM
CC       CYSTEINE TO PRODUCE ALANINE. SEEMS TO PARTICIPATE IN THE
CC       BIOSYNTHESIS OF THE NITROGENASE METALLOCLUSTERS BY PROVIDING THE
CC       INORGANIC SULFUR REQUIRED FOR THE FE-S CORE FORMATION (BY
CC       SIMILARITY).
CC   -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- MISCELLANEOUS: BELONGS TO THE NIF1 GENE CLUSTER WHICH IS EXPRESSED
CC       IN HETEROCYSTS UNDER ANAEROBIC AND AEROBIC CONDITIONS.
CC   -!- SIMILARITY: BELONGS TO CLASS-V OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC       AMINOTRANSFERASES. NIFS/ISCS SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X69898; CAA49523.1; -.
DR   InterPro; IPR000192; Aminotransf_class_V.
DR   PROSITE; PS00595; AA_TRANSFER_CLASS_5; PARTIAL.
KW   Nitrogen fixation; Lyase; Pyridoxal phosphate.
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1684 MW;  08B8F106DE65547D CRC64;


  Query Match             14.9%;  Score 24;  DB 1;  Length 15;
  Best Local Similarity   54.5%;  Pred. No. 1.3e+03;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    2;  Gaps    1;

Qy      12 ENNATFYFKID 22
           :||||   |:|
Db       7 DNNAT--TKVD 15


RESULT  16
FA71_TETPY
ID   FA71_TETPY     STANDARD;      PRT;    21 AA.
AC   P81424;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   71 kDa F-actin binding protein (Fragment).
OS   Tetrahymena pyriformis.
OC   Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;
OC   Tetrahymenina; Tetrahymena.
OX   NCBI_TaxID=5908;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=98207047; PubMed=9538250;
RA   Watanabe A., Kurasawa Y., Watanabe Y., Numata O.;
RT   "A new Tetrahymena actin-binding protein is localized in the division
RT   furrow.";
RL   J. Biochem. 123:607-613(1998).
CC   -!- FUNCTION: BINDS DIRECTLY TO F-ACTIN AND INDUCES ACTIN FILAMENT
CC       BUNDLING. MAY FUNCTION AS A REGULATOR OF ACTIN FILAMENT
CC       ORGANIZATION.
CC   -!- DEVELOPMENTAL STAGE: CO-LOCALIZED WITH ACTIN IN THE ORAL APPARATUS
CC       IN INTERPHASE CELLS. IN DIVIDING CELLS CO-LOCALIZED WITH ACTIN IN
CC       THE DIVISION FURROW.
CC   -!- PTM: THE N-TERMINUS IS BLOCKED.
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN IS:
CC       7.5, ITS MW IS: 71 kDa.
CC   -!- SIMILARITY: TO YEAST FIMBRIN.
KW   Actin-binding.
FT   NON_TER      21     21
SQ   SEQUENCE   21 AA;  2691 MW;  104FD6FD5E08FD28 CRC64;


  Query Match             14.9%;  Score 24;  DB 1;  Length 21;
  Best Local Similarity   80.0%;  Pred. No. 1.8e+03;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      18 YFKID 22
           :||||
Db       2 FFKID 6


RESULT  17
LPT_ECOLI
ID   LPT_ECOLI      STANDARD;      PRT;    21 AA.
AC   P03059;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Thr operon leader peptide (Thr operon attenuator).
GN   THRL OR B0001.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=79201669; PubMed=287010;
RA   Gardner J.F.;
RT   "Regulation of the threonine operon: tandem threonine and isoleucine
RT   codons in the control region and translational control of
RT   transcription termination.";
RL   Proc. Natl. Acad. Sci. U.S.A. 76:1706-1710(1979).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85264808; PubMed=2410621;
RA   Lynn S.P., Bauer C.E., Chapman K.A., Gardner J.F.;
RT   "Identification and characterization of mutants affecting
RT   transcription termination at the threonine operon attenuator.";
RL   J. Mol. Biol. 183:529-541(1985).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=95334362; PubMed=7610040;
RA   Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
RA   Blattner F.R.;
RT   "Analysis of the Escherichia coli genome VI: DNA sequence of the
```

```
RT   region from 92.8 through 100 minutes.";
RL   Nucleic Acids Res. 23:2105-2119(1995).
CC   -!- FUNCTION: THIS PROTEIN IS INVOLVED IN CONTROL OF THE BIOSYNTHESIS
CC       OF THREONINE.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V00360; CAA23658.1; -.
DR   EMBL; X68872; CAA48733.1; -.
DR   EMBL; J01706; AAA83913.1; -.
DR   EMBL; M28570; AAA24672.1; -.
DR   EMBL; U14003; AAA97300.1; -.
DR   EMBL; AE000111; AAC73112.1; -.
DR   PIR; A03595; LFECT.
DR   EcoGene; EG11277; thrL.
KW   Threonine biosynthesis; Leader peptide; Complete proteome.
SQ   SEQUENCE   21 AA;  2138 MW;  6A9C9167F89EECE4 CRC64;


  Query Match             14.9%;  Score 24;  DB 1;  Length 21;
  Best Local Similarity   71.4%;  Pred. No. 1.8e+03;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TLTHTIT 7
           |:| |||
Db       7 TITTTIT 13


RESULT  18
LPTN_ECOLI
ID   LPTN_ECOLI     STANDARD;      PRT;    24 AA.
AC   P09408;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Tryptophanase leader peptide.
GN   TNAL OR TNAC OR B3707 OR Z5202 OR ECS4644.
OS   Escherichia coli, and
OS   Escherichia coli O157:H7.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562, 83334;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=86033634; PubMed=3902796;
RA   Stewart V., Yanofsky C.;
RT   "Evidence for transcription antitermination control of tryptophanase
RT   operon expression in Escherichia coli K-12.";
RL   J. Bacteriol. 164:731-740(1985).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=93315143; PubMed=7686882;
RA   Burland V.D., Plunkett G. III, Daniels D.L., Blattner F.R.;
RT   "DNA sequence and analysis of 136 kilobases of the Escherichia coli
RT   genome: organizational symmetry around the origin of replication.";
RL   Genomics 16:551-561(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX   MEDLINE=21074935; PubMed=11206551;
RA   Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA   Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA   Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA   Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA   Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA   Welch R.A., Blattner F.R.;
RT   "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL   Nature 409:529-533(2001).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / RIMD 0509952;
RX   MEDLINE=21156231; PubMed=11258796;
RA   Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA   Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA   Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA   Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT   "Complete genome sequence of enterohemorrhagic Escherichia coli
RT   O157:H7 and genomic comparison with a laboratory strain K-12.";
RL   DNA Res. 8:11-22(2001).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M11990; AAA24678.1; -.
DR   EMBL; L10328; AAA62058.1; -.
DR   EMBL; AE000448; AAC76730.1; -.
DR   EMBL; AE005602; AAG58907.1; -.
DR   EMBL; AP002566; BAB38067.1; -.
DR   EcoGene; EG11276; tnaL.
KW   Leader peptide; Complete proteome.
SQ   SEQUENCE   24 AA;  2894 MW;  66E3987EA7C052F9 CRC64;


  Query Match             14.9%;  Score 24;  DB 1;  Length 24;
  Best Local Similarity   66.7%;  Pred. No. 2.1e+03;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      18 YFKIDN 23
           :| |||
Db      12 WFNIDN 17


RESULT  19
PUTA_KLEPN
ID   PUTA_KLEPN     STANDARD;      PRT;    26 AA.
AC   P23725;
DT   01-NOV-1991 (Rel. 20, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Bifunctional putA protein [Includes: Proline dehydrogenase
DE   (EC 1.5.99.8) (Proline oxidase); Delta-1-pyrroline-5-carboxylate
DE   dehydrogenase (EC 1.5.1.12) (P5C dehydrogenase)] (Fragment).
GN   PUTA.
OS   Klebsiella pneumoniae.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Klebsiella.
OX   NCBI_TaxID=573;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91100369; PubMed=1987164;
RA   Chen L.M., Maloy S.;
RT   "Regulation of proline utilization in enteric bacteria: cloning and
RT   characterization of the Klebsiella put control region.";
RL   J. Bacteriol. 173:783-790(1991).
CC   -!- FUNCTION: OXIDIZES PROLINE TO GLUTAMATE FOR USE AS A CARBON AND
CC       NITROGEN SOURCE AND ALSO FUNCTION AS A TRANSCRIPTIONAL REPRESSOR
CC       OF THE PUT OPERON.
CC   -!- CATALYTIC ACTIVITY: L-proline + acceptor + H(2)O = (S)-1-
CC       pyrroline-5-carboxylate + reduced acceptor.
CC   -!- COFACTOR: FAD.
CC   -!- PATHWAY: PROLINE UTILIZATION.
CC   -!- INDUCTION: BY PROLINE, AUTOREPRESSION AND CATABOLITE REPRESSION,
CC       AND IS POTENTIALLY NITROGEN CONTROLLED.
```

```
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M63160; AAA25139.1; -.
DR   InterPro; IPR002086; Aldehyde_dehydr.
DR   PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; PARTIAL.
DR   PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; PARTIAL.
KW   Multifunctional enzyme; Oxidoreductase; Flavoprotein; FAD; NAD;
KW   Transcription regulation; Repressor; DNA-binding; Proline metabolism.
FT   NON_TER      26     26
SQ   SEQUENCE   26 AA;  2824 MW;  BB332D0DE504CE19 CRC64;


  Query Match             14.9%;  Score 24;  DB 1;  Length 26;
  Best Local Similarity   26.7%;  Pred. No. 2.3e+03;
  Matches    4;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy     16 TFYFKIDNVKKARVQ 30
          |   |:|:  : |::
Db      5 TMGVKLDDATRERIK 19


RESULT  20
AMPT_BACST
ID   AMPT_BACST     STANDARD;      PRT;    30 AA.
AC   P00728;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Thermophilic aminopeptidase I alpha chain (EC 3.4.11.-) (Fragment).
OS   Bacillus stearothermophilus.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Geobacillus.
OX   NCBI_TaxID=1422;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=NCIB 8924;
RX   MEDLINE=74087636; PubMed=4521203;
RA   Stoll E., Ericsson L.H., Zuber H.;
RT   "The function of the two subunits of thermophilic aminopeptidase I.";
RL   Proc. Natl. Acad. Sci. U.S.A. 70:3781-3784(1973).
CC   -!- FUNCTION: METALLOENZYME OF HIGH TEMPERATURE STABILITY AND OF
CC       BROAD SPECIFICITY, RELEASING ALL N-TERMINAL AMINO ACIDS.
CC   -!- SUBUNIT: 12 CHAINS OF TWO DIFFERENT BUT HOMOLOGOUS TYPES, ALPHA
CC       AND BETA, WHICH CAN COMBINE IN VARIOUS RATIOS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M42.
DR   PIR; A00908; AIBSAF.
DR   MEROPS; M42.002; -.
KW   Hydrolase; Aminopeptidase; Metalloprotease.
FT   NON_TER      30     30
SQ   SEQUENCE   30 AA;  3274 MW;  D712C9C23E618142 CRC64;


  Query Match             14.9%;  Score 24;  DB 1;  Length 30;
  Best Local Similarity   42.9%;  Pred. No. 2.6e+03;
  Matches    6;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      2 LTHTITKLNAENNA 15
          |  |:| | |  :|
Db      3 LDETLTMLKALTDA 16


Search completed: July  8, 2002, 11:53:13
Job time : 199 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 8, 2002, 11:49:19 ; Search time 23.81 Seconds
(without alignments)
232.501 Million cell updates/sec

Title: US-09-461-061A-4
Perfect score: 161
Sequence: 1 TLTHTITKLNAENNATFYFKIDNVKKARVQVV 32

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 14647

Minimum DB seq length: 0
Maximum DB seq length: 32

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	33	20.5	13	4	Q9UC27	Q9uc27 homo sapien
2	32	19.9	27	12	Q9QH83	Q9qh83 hepatitis c
3	32	19.9	27	12	Q9QH82	Q9qh82 hepatitis c
4	32	19.9	27	12	Q9QH81	Q9qh81 hepatitis c
5	32	19.9	27	12	Q9QH80	Q9qh80 hepatitis c
6	32	19.9	27	12	Q9QH79	Q9qh79 hepatitis c
7	32	19.9	27	12	Q9QH78	Q9qh78 hepatitis c
8	32	19.9	27	12	Q9QH77	Q9qh77 hepatitis c
9	32	19.9	27	12	Q9QH76	Q9qh76 hepatitis c
10	32	19.9	27	12	Q9QH75	Q9qh75 hepatitis c
11	32	19.9	27	12	Q9QH74	Q9qh74 hepatitis c
12	32	19.9	27	12	Q9QH73	Q9qh73 hepatitis c
13	32	19.9	27	12	Q9QH72	Q9qh72 hepatitis c
14	32	19.9	27	12	Q9QH70	Q9qh70 hepatitis c
15	32	19.9	27	12	Q9QH69	Q9qh69 hepatitis c
16	32	19.9	27	12	Q9QH68	Q9qh68 hepatitis c
17	32	19.9	27	12	Q9QH66	Q9qh66 hepatitis c
18	32	19.9	32	6	Q9TR69	Q9tr69 sus scrofa
19	32	19.9	32	6	Q9TR67	Q9tr67 sus scrofa
20	31	19.3	23	5	O44207	O44207 onchocerca
21	31	19.3	30	2	Q49048	Q49048 mycoplasma
22	31	19.3	32	13	Q9W7P3	Q9w7p3 morone saxa
23	30.5	18.9	32	12	Q914F9	Q914f9 sulfolobus
24	30	18.6	23	4	Q96IP0	Q96ip0 homo sapien
25	30	18.6	25	6	Q95L28	Q95128 canis famil
26	30	18.6	25	16	O26056	O26056 helicobacte
27	30	18.6	26	11	Q9CTD3	Q9ctd3 mus musculu
28	30	18.6	30	4	Q96QY6	Q96qy6 homo sapien
29	30	18.6	31	4	Q9BQU1	Q9bqu1 homo sapien
30	29	18.0	20	13	Q9PSI5	Q9psi5 oncorhynchu
31	29	18.0	23	5	Q9XZW5	Q9xzw5 littorina l
32	29	18.0	23	5	Q9Y003	Q9y003 melarhaphe
33	28.5	17.7	29	13	Q90817	Q90817 gallus gall
34	28	17.4	14	5	Q26100	Q26100 pratylenchu
35	28	17.4	16	5	Q26101	Q26101 pratylenchu
36	28	17.4	19	5	Q26099	Q26099 pratylenchu
37	28	17.4	21	2	Q9R585	Q9r585 vibrio. end
38	28	17.4	23	5	Q9XZZ7	Q9xzz7 littorina s
39	28	17.4	23	5	Q9XZW1	Q9xzw1 littorina a
40	28	17.4	23	5	Q9XZW4	Q9xzw4 littorina l
41	28	17.4	26	4	Q9UC79	Q9uc79 homo sapien
42	28	17.4	27	12	Q9QRS2	Q9qrs2 hepatitis c
43	28	17.4	27	12	Q9QI47	Q9qi47 hepatitis c
44	28	17.4	27	12	Q9QI44	Q9qi44 hepatitis c
45	28	17.4	27	12	Q9QI41	Q9qi41 hepatitis c

ALIGNMENTS

RESULT 1
Q9UC27
ID Q9UC27 PRELIMINARY; PRT; 13 AA.
AC Q9UC27;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE 94 KDA EPIDIDYMAL CYTOKERATIN-LIKE PROTEIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=95226590; PubMed=7536047;
RA Boue F., Duquenne C., Lassalle B., Lefevre A., Finaz C.;
RT "FLB1, a human protein of epididymal origin that is involved in the
RT sperm-oocyte recognition process.";
RL Biol. Reprod. 52:267-278(1995).
SQ SEQUENCE 13 AA; 1610 MW; 000A683B81B7A45B CRC64;

Query Match 20.5%; Score 33; DB 4; Length 13;
Best Local Similarity 46.2%; Pred. No. 4.2e+02;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

```
Qy      13 NNATFYFKIDNVK 25
           | :|:|  ||::|
Db       1 NYSTYYNTIDDLK 13
```

RESULT 2
Q9QH83
ID Q9QH83 PRELIMINARY; PRT; 27 AA.
AC Q9QH83;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70)

```
DE   (NS1)] (FRAGMENT).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sandres K., Dubois M., Pasquier C., Izopet J.;
RT   "The genetic heterogeneity of the hypervariable region 1 of the viral
RT   genome and the sensitivity of hepatitis C virus to interferon alpha
RT   therapy.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF167045; AAD53678.1; -.
DR   InterPro; IPR002531; HCV_NS1.
DR   Pfam; PF01560; HCV_NS1; 1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
FT   NON_TER       1      1
FT   NON_TER      27     27
SQ   SEQUENCE   27 AA;  2754 MW;  D4385958667D5361 CRC64;


  Query Match             19.9%;  Score 32;  DB 12;  Length 27;
  Best Local Similarity   33.3%;  Pred. No. 1.3e+03;
  Matches    5;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      3 THTITKLNAENNATF 17
          |||:  :::   |:|
Db      2 THTVGSISSRATASF 16


RESULT   3
Q9QH82
ID   Q9QH82      PRELIMINARY;      PRT;    27 AA.
AC   Q9QH82;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70)
DE   (NS1)] (FRAGMENT).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sandres K., Dubois M., Pasquier C., Izopet J.;
RT   "The genetic heterogeneity of the hypervariable region 1 of the viral
RT   genome and the sensitivity of hepatitis C virus to interferon alpha
RT   therapy.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF167046; AAD53679.1; -.
DR   InterPro; IPR002531; HCV_NS1.
DR   Pfam; PF01560; HCV_NS1; 1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
FT   NON_TER       1      1
FT   NON_TER      27     27
SQ   SEQUENCE   27 AA;  2754 MW;  D4385958667D5361 CRC64;


  Query Match             19.9%;  Score 32;  DB 12;  Length 27;
  Best Local Similarity   33.3%;  Pred. No. 1.3e+03;
  Matches    5;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      3 THTITKLNAENNATF 17
          |||:  :::   |:|
Db      2 THTVGSISSRATASF 16


RESULT   4
Q9QH81
ID   Q9QH81      PRELIMINARY;      PRT;    27 AA.
AC   Q9QH81;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70)
DE   (NS1)] (FRAGMENT).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sandres K., Dubois M., Pasquier C., Izopet J.;
RT   "The genetic heterogeneity of the hypervariable region 1 of the viral
RT   genome and the sensitivity of hepatitis C virus to interferon alpha
RT   therapy.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF167047; AAD53680.1; -.
DR   InterPro; IPR002531; HCV_NS1.
DR   Pfam; PF01560; HCV_NS1; 1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
FT   NON_TER       1      1
FT   NON_TER      27     27
SQ   SEQUENCE   27 AA;  2754 MW;  D4385958667D5361 CRC64;


  Query Match             19.9%;  Score 32;  DB 12;  Length 27;
  Best Local Similarity   33.3%;  Pred. No. 1.3e+03;
  Matches    5;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      3 THTITKLNAENNATF 17
          |||:  :::   |:|
Db      2 THTVGSISSRATASF 16


RESULT   5
Q9QH80
ID   Q9QH80      PRELIMINARY;      PRT;    27 AA.
AC   Q9QH80;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70)
DE   (NS1)] (FRAGMENT).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sandres K., Dubois M., Pasquier C., Izopet J.;
RT   "The genetic heterogeneity of the hypervariable region 1 of the viral
RT   genome and the sensitivity of hepatitis C virus to interferon alpha
RT   therapy.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF167048; AAD53681.1; -.
DR   InterPro; IPR002531; HCV_NS1.
DR   Pfam; PF01560; HCV_NS1; 1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
FT   NON_TER       1      1
FT   NON_TER      27     27
SQ   SEQUENCE   27 AA;  2754 MW;  D4385958667D5361 CRC64;


  Query Match             19.9%;  Score 32;  DB 12;  Length 27;
  Best Local Similarity   33.3%;  Pred. No. 1.3e+03;
  Matches    5;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      3 THTITKLNAENNATF 17
          |||:  :::   |:|
Db      2 THTVGSISSRATASF 16
```

```
RESULT   6
Q9QH79
ID   Q9QH79      PRELIMINARY;      PRT;    27 AA.
AC   Q9QH79;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70)
DE   (NS1)] (FRAGMENT).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sandres K., Dubois M., Pasquier C., Izopet J.;
RT   "The genetic heterogeneity of the hypervariable region 1 of the viral
RT   genome and the sensitivity of hepatitis C virus to interferon alpha
RT   therapy.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF167049; AAD53682.1; -.
DR   InterPro; IPR002531; HCV_NS1.
DR   Pfam; PF01560; HCV_NS1; 1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
FT   NON_TER       1      1
FT   NON_TER      27     27
SQ   SEQUENCE   27 AA;  2754 MW;  D4385958667D5361 CRC64;


  Query Match             19.9%;  Score 32;  DB 12;  Length 27;
  Best Local Similarity   33.3%;  Pred. No. 1.3e+03;
  Matches    5;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      3 THTITKLNAENNATF 17
          |||:  :::   |:|
Db      2 THTVGSISSRATASF 16


RESULT   7
Q9QH78
ID   Q9QH78      PRELIMINARY;      PRT;    27 AA.
AC   Q9QH78;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70)
DE   (NS1)] (FRAGMENT).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sandres K., Dubois M., Pasquier C., Izopet J.;
RT   "The genetic heterogeneity of the hypervariable region 1 of the viral
RT   genome and the sensitivity of hepatitis C virus to interferon alpha
RT   therapy.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF167050; AAD53683.1; -.
DR   InterPro; IPR002531; HCV_NS1.
DR   Pfam; PF01560; HCV_NS1; 1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
FT   NON_TER       1      1
FT   NON_TER      27     27
SQ   SEQUENCE   27 AA;  2754 MW;  D4385958667D5361 CRC64;


  Query Match             19.9%;  Score 32;  DB 12;  Length 27;
  Best Local Similarity   33.3%;  Pred. No. 1.3e+03;
  Matches    5;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      3 THTITKLNAENNATF 17
          |||:  :::   |:|
Db      2 THTVGSISSRATASF 16


RESULT   8
Q9QH77
ID   Q9QH77      PRELIMINARY;      PRT;    27 AA.
AC   Q9QH77;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70)
DE   (NS1)] (FRAGMENT).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sandres K., Dubois M., Pasquier C., Izopet J.;
RT   "The genetic heterogeneity of the hypervariable region 1 of the viral
RT   genome and the sensitivity of hepatitis C virus to interferon alpha
RT   therapy.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF167051; AAD53684.1; -.
DR   InterPro; IPR002531; HCV_NS1.
DR   Pfam; PF01560; HCV_NS1; 1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
FT   NON_TER       1      1
FT   NON_TER      27     27
SQ   SEQUENCE   27 AA;  2754 MW;  D4385958667D5361 CRC64;


  Query Match             19.9%;  Score 32;  DB 12;  Length 27;
  Best Local Similarity   33.3%;  Pred. No. 1.3e+03;
  Matches    5;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      3 THTITKLNAENNATF 17
          |||:  :::   |:|
Db      2 THTVGSISSRATASF 16


RESULT   9
Q9QH76
ID   Q9QH76      PRELIMINARY;      PRT;    27 AA.
AC   Q9QH76;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70)
DE   (NS1)] (FRAGMENT).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sandres K., Dubois M., Pasquier C., Izopet J.;
RT   "The genetic heterogeneity of the hypervariable region 1 of the viral
RT   genome and the sensitivity of hepatitis C virus to interferon alpha
RT   therapy.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF167052; AAD53685.1; -.
DR   InterPro; IPR002531; HCV_NS1.
DR   Pfam; PF01560; HCV_NS1; 1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
FT   NON_TER       1      1
FT   NON_TER      27     27
```

```
SQ   SEQUENCE   27 AA;  2754 MW;  D4385958667D5361 CRC64;


  Query Match             19.9%;  Score 32;  DB 12;  Length 27;
  Best Local Similarity   33.3%;  Pred. No. 1.3e+03;
  Matches    5;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      3 THTITKLNAENNATF 17
          |||:  :::   |:|
Db      2 THTVGSISSRATASF 16


RESULT  10
Q9QH75
ID   Q9QH75      PRELIMINARY;      PRT;    27 AA.
AC   Q9QH75;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70)
DE   (NS1)] (FRAGMENT).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sandres K., Dubois M., Pasquier C., Izopet J.;
RT   "The genetic heterogeneity of the hypervariable region 1 of the viral
RT   genome and the sensitivity of hepatitis C virus to interferon alpha
RT   therapy.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF167053; AAD53686.1; -.
DR   InterPro; IPR002531; HCV_NS1.
DR   Pfam; PF01560; HCV_NS1; 1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
FT   NON_TER       1      1
FT   NON_TER      27     27
SQ   SEQUENCE   27 AA;  2754 MW;  D4385958667D5361 CRC64;


  Query Match             19.9%;  Score 32;  DB 12;  Length 27;
  Best Local Similarity   33.3%;  Pred. No. 1.3e+03;
  Matches    5;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      3 THTITKLNAENNATF 17
          |||:  :::   |:|
Db      2 THTVGSISSRATASF 16


RESULT  11
Q9QH74
ID   Q9QH74      PRELIMINARY;      PRT;    27 AA.
AC   Q9QH74;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70)
DE   (NS1)] (FRAGMENT).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sandres K., Dubois M., Pasquier C., Izopet J.;
RT   "The genetic heterogeneity of the hypervariable region 1 of the viral
RT   genome and the sensitivity of hepatitis C virus to interferon alpha
RT   therapy.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF167054; AAD53687.1; -.
DR   InterPro; IPR002531; HCV_NS1.
DR   Pfam; PF01560; HCV_NS1; 1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
FT   NON_TER       1      1
FT   NON_TER      27     27
SQ   SEQUENCE   27 AA;  2740 MW;  D43859586668B361 CRC64;


  Query Match             19.9%;  Score 32;  DB 12;  Length 27;
  Best Local Similarity   33.3%;  Pred. No. 1.3e+03;
  Matches    5;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      3 THTITKLNAENNATF 17
          |||:  :::   |:|
Db      2 THTVGSISSRATASF 16


RESULT  12
Q9QH73
ID   Q9QH73      PRELIMINARY;      PRT;    27 AA.
AC   Q9QH73;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70)
DE   (NS1)] (FRAGMENT).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sandres K., Dubois M., Pasquier C., Izopet J.;
RT   "The genetic heterogeneity of the hypervariable region 1 of the viral
RT   genome and the sensitivity of hepatitis C virus to interferon alpha
RT   therapy.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF167055; AAD53688.1; -.
DR   InterPro; IPR002531; HCV_NS1.
DR   Pfam; PF01560; HCV_NS1; 1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
FT   NON_TER       1      1
FT   NON_TER      27     27
SQ   SEQUENCE   27 AA;  2754 MW;  D4385958667D5361 CRC64;


  Query Match             19.9%;  Score 32;  DB 12;  Length 27;
  Best Local Similarity   33.3%;  Pred. No. 1.3e+03;
  Matches    5;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      3 THTITKLNAENNATF 17
          |||:  :::   |:|
Db      2 THTVGSISSRATASF 16


RESULT  13
Q9QH72
ID   Q9QH72      PRELIMINARY;      PRT;    27 AA.
AC   Q9QH72;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70)
DE   (NS1)] (FRAGMENT).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sandres K., Dubois M., Pasquier C., Izopet J.;
RT   "The genetic heterogeneity of the hypervariable region 1 of the viral
```

```
RT   genome and the sensitivity of hepatitis C virus to interferon alpha
RT   therapy.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF167056; AAD53689.1; -.
DR   InterPro; IPR002531; HCV_NS1.
DR   Pfam; PF01560; HCV_NS1; 1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
FT   NON_TER       1      1
FT   NON_TER      27     27
SQ   SEQUENCE   27 AA;  2754 MW;  D4385958667D5361 CRC64;


  Query Match             19.9%;  Score 32;  DB 12;  Length 27;
  Best Local Similarity   33.3%;  Pred. No. 1.3e+03;
  Matches    5;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      3 THTITKLNAENNATF 17
          |||:  :::   |:|
Db      2 THTVGSISSRATASF 16


RESULT  14
Q9QH70
ID   Q9QH70      PRELIMINARY;      PRT;    27 AA.
AC   Q9QH70;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70)
DE   (NS1)] (FRAGMENT).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sandres K., Dubois M., Pasquier C., Izopet J.;
RT   "The genetic heterogeneity of the hypervariable region 1 of the viral
RT   genome and the sensitivity of hepatitis C virus to interferon alpha
RT   therapy.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF167058; AAD53691.1; -.
DR   InterPro; IPR002531; HCV_NS1.
DR   Pfam; PF01560; HCV_NS1; 1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
FT   NON_TER       1      1
FT   NON_TER      27     27
SQ   SEQUENCE   27 AA;  2754 MW;  D4385958667D5361 CRC64;


  Query Match             19.9%;  Score 32;  DB 12;  Length 27;
  Best Local Similarity   33.3%;  Pred. No. 1.3e+03;
  Matches    5;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      3 THTITKLNAENNATF 17
          |||:  :::   |:|
Db      2 THTVGSISSRATASF 16


RESULT  15
Q9QH69
ID   Q9QH69      PRELIMINARY;      PRT;    27 AA.
AC   Q9QH69;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70)
DE   (NS1)] (FRAGMENT).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sandres K., Dubois M., Pasquier C., Izopet J.;
RT   "The genetic heterogeneity of the hypervariable region 1 of the viral
RT   genome and the sensitivity of hepatitis C virus to interferon alpha
RT   therapy.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF167059; AAD53692.1; -.
DR   InterPro; IPR002531; HCV_NS1.
DR   Pfam; PF01560; HCV_NS1; 1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
FT   NON_TER       1      1
FT   NON_TER      27     27
SQ   SEQUENCE   27 AA;  2754 MW;  D4385958667D5361 CRC64;


  Query Match             19.9%;  Score 32;  DB 12;  Length 27;
  Best Local Similarity   33.3%;  Pred. No. 1.3e+03;
  Matches    5;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      3 THTITKLNAENNATF 17
          |||:  :::   |:|
Db      2 THTVGSISSRATASF 16


RESULT  16
Q9QH68
ID   Q9QH68      PRELIMINARY;      PRT;    27 AA.
AC   Q9QH68;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70)
DE   (NS1)] (FRAGMENT).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sandres K., Dubois M., Pasquier C., Izopet J.;
RT   "The genetic heterogeneity of the hypervariable region 1 of the viral
RT   genome and the sensitivity of hepatitis C virus to interferon alpha
RT   therapy.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF167060; AAD53693.1; -.
DR   InterPro; IPR002531; HCV_NS1.
DR   Pfam; PF01560; HCV_NS1; 1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
FT   NON_TER       1      1
FT   NON_TER      27     27
SQ   SEQUENCE   27 AA;  2754 MW;  D4385958667D5361 CRC64;


  Query Match             19.9%;  Score 32;  DB 12;  Length 27;
  Best Local Similarity   33.3%;  Pred. No. 1.3e+03;
  Matches    5;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      3 THTITKLNAENNATF 17
          |||:  :::   |:|
Db      2 THTVGSISSRATASF 16


RESULT  17
Q9QH66
ID   Q9QH66      PRELIMINARY;      PRT;    27 AA.
AC   Q9QH66;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
```

```
DE   GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70)
DE   (NS1)] (FRAGMENT).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sandres K., Dubois M., Pasquier C., Izopet J.;
RT   "The genetic heterogeneity of the hypervariable region 1 of the viral
RT   genome and the sensitivity of hepatitis C virus to interferon alpha
RT   therapy.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF167062; AAD53695.1; -.
DR   InterPro; IPR002531; HCV_NS1.
DR   Pfam; PF01560; HCV_NS1; 1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
FT   NON_TER       1      1
FT   NON_TER      27     27
SQ   SEQUENCE   27 AA;  2754 MW;  D4385958667D5361 CRC64;


  Query Match             19.9%;  Score 32;  DB 12;  Length 27;
  Best Local Similarity   33.3%;  Pred. No. 1.3e+03;
  Matches    5;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 THTITKLNAENNATF 17
           |||:  :::   |:|
Db       2 THTVGSISSRATASF 16


RESULT  18
Q9TR69
ID   Q9TR69      PRELIMINARY;      PRT;    32 AA.
AC   Q9TR69;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   PI2 I ALPHA-PROTEASE INHIBITOR (FRAGMENT).
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=95268713; PubMed=7749636;
RA   Stratil A., Cizova-Schroffelova D., Gabrisova E., Pavlik M.,
RA   Coppieters W., Peelman L., Van de Weghe A., Bouquet Y.;
RT   "Pig plasma alpha-protease inhibitors PI2, PI3 and PI4 are members of
RT   the antichymotrypsin family.";
RL   Comp. Biochem. Physiol. 111B:53-60(1995).
FT   NON_TER       1      1
FT   NON_TER      32     32
SQ   SEQUENCE   32 AA;  3442 MW;  467B14B3FB6E72DA CRC64;


  Query Match             19.9%;  Score 32;  DB 6;  Length 32;
  Best Local Similarity   53.3%;  Pred. No. 1.5e+03;
  Matches    8;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKLNAENNA 15
           ||   ||||   | |
Db      10 TLKDQITKLPVHNTA 24


RESULT  19
Q9TR67
ID   Q9TR67      PRELIMINARY;      PRT;    32 AA.
AC   Q9TR67;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   PI4 C2 ALPHA-PROTEASE INHIBITOR (FRAGMENT).
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=95268713; PubMed=7749636;
RA   Stratil A., Cizova-Schroffelova D., Gabrisova E., Pavlik M.,
RA   Coppieters W., Peelman L., Van de Weghe A., Bouquet Y.;
RT   "Pig plasma alpha-protease inhibitors PI2, PI3 and PI4 are members of
RT   the antichymotrypsin family.";
RL   Comp. Biochem. Physiol. 111B:53-60(1995).
FT   NON_TER       1      1
FT   NON_TER      32     32
SQ   SEQUENCE   32 AA;  3464 MW;  4CC602E86D23B2DA CRC64;


  Query Match             19.9%;  Score 32;  DB 6;  Length 32;
  Best Local Similarity   40.9%;  Pred. No. 1.5e+03;
  Matches    9;  Conservative    1;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 TLTHTITKLNAENNATFYFKID 22
           ||   | || | | |  :   |
Db      10 TLKGQIKKLPAVNTAVVWXNXD 31


RESULT  20
O44207
ID   O44207      PRELIMINARY;      PRT;    23 AA.
AC   O44207;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ANTIGEN (FRAGMENT).
OS   Onchocerca volvulus.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC   Onchocercidae; Onchocerca.
OX   NCBI_TaxID=6282;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FOREST;
RA   Titanji V.P.K., Sakwe A.M., Ghogomu S.M., Souopgui J., Djokam R.T.,
RA   Perler F., Rask L.;
RT   "A cDNA coding for a novel antigen from Onchocerca volvulus.";
RL   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Y15882; CAA75807.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   23 AA;  2777 MW;  CA40C0430F5F1F54 CRC64;


  Query Match             19.3%;  Score 31;  DB 5;  Length 23;
  Best Local Similarity   35.3%;  Pred. No. 1.5e+03;
  Matches    6;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       4 HTITKLNAENNATFYFK 20
           |    |:  :|| :| |
Db       4 HINDYLDTNHNADYYLK 20


Search completed: July  8, 2002, 11:52:56
Job time: 217 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 8, 2002, 11:47:19 ; Search time 29.33 Seconds
(without alignments)
121.185 Million cell updates/sec

Title: US-09-461-061A-4
Perfect score: 161
Sequence: 1 TLTHTITKLNAENNATFYFKIDNVKKARVQVV 32

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 285274

Minimum DB seq length: 0
Maximum DB seq length: 32

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	161	100.0	32	21	AAY95408	Anti-angiogenic D3
2	95	59.0	26	18	AAW54336	Bradykinin analogo
3	84	52.2	16	21	AAY95410	Anti-angiogenic D3
4	83	51.6	16	21	AAY95409	Anti-angiogenic D3
5	74	46.0	16	21	AAB08553	Peptide identified
6	60	37.3	12	21	AAY95406	Anti-angiogenic pe
7	56	34.8	12	21	AAY95407	Anti-angiogenic pe
8	49	30.4	9	21	AAB37455	Human kininogen D3
9	45	28.0	10	21	AAY95405	Anti-angiogenic D3
10	40	24.8	32	20	AAW94743	Anti-Staph (HAY) 9
11	39	24.2	32	13	AAR27043	Human light chain
12	39	24.2	32	16	AAR85155	Human ONS-M21 anti
13	39	24.2	32	16	AAR77206	Mouse anti-human I
14	39	24.2	32	17	AAR96288	Light chain framew
15	39	24.2	32	17	AAR97321	Humanised monoclon
16	37	23.0	32	17	AAR87043	Human group I ligh
17	37	23.0	32	20	AAY52745	Humanised ATR-5 L
18	37	23.0	32	20	AAY52746	Humanised ATR-5 L
19	37	23.0	32	20	AAY52747	Humanised ATR-5 L
20	37	23.0	32	22	AAB98286	Anti-A33 antigen i
21	37	23.0	32	22	AAB98287	Anti-A33 antigen i
22	37	23.0	32	22	AAB98288	Anti-A33 antigen i
23	37	23.0	32	22	AAB97666	A33 antigen bindin
24	37	23.0	32	22	AAB97667	A33 antigen bindin
25	37	23.0	32	22	AAB97668	A33 antigen bindin
26	36.5	22.7	24	22	AAO04947	Human polypeptide
27	36	22.4	18	15	AAR47004	Ig kappa chain pos
28	35	21.7	7	18	AAW54340	Bradykinin analogo
29	35	21.7	7	19	AAW77428	Peptide LNA7, a no
30	35	21.7	22	18	AAW17883	Photorhabdus lumin
31	35	21.7	22	19	AAW56565	Toxin fragment of
32	35	21.7	27	12	AAR10432	CD4 fragment pepti
33	35	21.7	27	14	AAR31344	HIV superinfection
34	35	21.7	32	22	AAB98290	Anti-A33 antigen i
35	35	21.7	32	22	AAB97670	A33 antigen bindin
36	34	21.1	21	21	AAY94747	Non-maltogenic exo
37	34	21.1	28	21	AAY89463	Core polypeptide f
38	34	21.1	28	22	ABG18695	Novel human diagno
39	34	21.1	28	22	ABB00871	Viral DP178/107-li
40	34	21.1	28	22	ABB02346	Viral core polypep
41	34	21.1	28	22	AAU13417	DP178-like/DP107-l
42	34	21.1	28	22	AAB77864	Core polypeptide T
43	34	21.1	30	19	AAW73005	Cobra venom mocarh
44	34	21.1	31	20	AAY30773	Vasoactive intesti
45	34	21.1	32	17	AAR97330	Humanised monoclon

ALIGNMENTS

RESULT 1
AAY95408
ID AAY95408 standard; Peptide; 32 AA.
XX
AC AAY95408;
XX
DT 25-SEP-2000 (first entry)
XX
DE Anti-angiogenic D3 peptide.
XX
KW Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;
KW endothelial cell proliferation; apoptosis; cancer; ocular disorder;
KW rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic;
KW therapy; human; D3 peptide.
XX
OS Homo sapiens.
XX
PN WO200035407-A2.
XX
PD 22-JUN-2000.
XX
PF 02-DEC-1999; 99WO-US28465.
XX
PR 16-DEC-1998; 98US-0112427.
XX
PA (UTEM) UNIV TEMPLE.
PA (MCCR/) MCCRAE R K.
XX
PI McCrae RK;
XX
DR WPI; 2000-442247/38.
XX
PT Composition for inhibiting angiogenesis and endothelial cell
PT proliferation, inducing endothelial cell apoptosis and treating cancer,

PT rheumatoid arthritis, and ocular disorders comprises a kininogen domain
PT 3 analog -
XX
PS Claim 4; Page 26; 44pp; English.
XX
CC The present sequence is that of a D3 peptide derived from human
CC high mol.wt. kininogen (HK) domain 3 (see AAY95426). The D3 peptide
CC inhibits endothelial cell proliferation and thus possesses
CC anti-angiogenic activity. It is an example of D3 peptides of the
CC invention (see AAY95405-26) that are analogues of certain sites in
CC the HK domain 3, in this case amino acid residues Asn275-Lys282.
CC The peptides inhibit endothelial cell proliferation and may also
CC induce endothelial cell apoptosis. Compositions including the
CC peptides are used in claimed methods for inhibiting angiogenesis,
CC inhibiting endothelial cell proliferation, and inducing endothelial
CC cell apoptosis. Cancer, rheumatoid arthritis, and ocular disorders
CC characterized by undesired vascularization of the retina are treated.
XX
SQ Sequence 32 AA;

```
  Query Match             100.0%;  Score 161;  DB 21;  Length 32;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-15;
  Matches   32;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 TLTHTITKLNAENNATFYFKIDNVKKARVQVV 32
          ||||||||||||||||||||||||||||||||
Db      1 tlthtitklnaennatfyfkidnvkkarvqvv 32
```

RESULT 2
AAW54336
ID AAW54336 standard; peptide; 26 AA.
XX
AC AAW54336;
XX
DT 30-JUL-1998 (first entry)
XX
DE Bradykinin analogous peptide 19.
XX
KW Inhibition; thrombin-induced platelet; prevention; platelet aggregation;
KW ADP-induced activation.
XX
OS Synthetic.
XX
PN WO9641640-A1.
XX
PD 27-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US09940.
XX
PR 09-JUN-1995; 95US-0000096.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Hasan AAK, Schmaier AH;
XX
DR WPI; 1997-065304/06.
XX
PT Inhibition of platelet activation and aggregation - by admin. of new
PT or known bradykinin analogues
XX
PS Disclosure; Page 44; 73pp; English.
XX
CC Administration of a peptide or multimer related to bradykinin or other
CC disclosed peptides and multimers can be used for the inhibition of
CC thrombin-induced platelets or other cells. They can also be used for
CC preventing platelet aggregation, or inhibiting ADP-induced activation.
CC This is useful to prevent arterial occlusions arising from coronary
CC thrombosis and stroke.
XX
SQ Sequence 26 AA;

```
  Query Match             59.0%;  Score 95;  DB 18;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-06;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      14 NATFYFKIDNVKKARVQVV 32
           ||||||||||||||||||
Db       1 natfyfkidnvkkarvqvv 19
```

RESULT 3
AAY95410
ID AAY95410 standard; Peptide; 16 AA.
XX
AC AAY95410;
XX
DT 25-SEP-2000 (first entry)
XX
DE Anti-angiogenic D3 peptide.
XX
KW Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;
KW endothelial cell proliferation; apoptosis; cancer; ocular disorder;
KW rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic;
KW therapy; human; D3 peptide.
XX
OS Homo sapiens.
XX
PN WO200035407-A2.
XX
PD 22-JUN-2000.
XX
PF 02-DEC-1999; 99WO-US28465.
XX
PR 16-DEC-1998; 98US-0112427.
XX
PA (UTEM) UNIV TEMPLE.
PA (MCCR/) MCCRAE R K.
XX
PI McCrae RK;
XX
DR WPI; 2000-442247/38.
XX
PT Composition for inhibiting angiogenesis and endothelial cell
PT proliferation, inducing endothelial cell apoptosis and treating cancer,
PT rheumatoid arthritis, and ocular disorders comprises a kininogen domain
PT 3 analog -
XX
PS Claim 7; Page 26; 44pp; English.
XX
CC The present sequence is that of a D3 peptide derived from human
CC high mol.wt. kininogen (HK) domain 3 (see AAY95426). The D3 peptide
CC inhibits endothelial cell proliferation and thus possesses
CC anti-angiogenic activity. It is an example of D3 peptides of the
CC invention (see AAY95405-26) that are analogues of certain sites in
CC the HK domain 3, in this case amino acid residues Asn275-Lys282.
CC The peptides inhibit endothelial cell proliferation and may also
CC induce endothelial cell apoptosis. Compositions including the
CC peptides are used in claimed methods for inhibiting angiogenesis,
CC inhibiting endothelial cell proliferation, and inducing endothelial
CC cell apoptosis. Cancer, rheumatoid arthritis, and ocular disorders
CC characterized by undesired vascularization of the retina are treated.
CC The IC50 value for the present peptide was less than 0.8 uM for
CC inhibition of fibroblast growth factor-induced HUVEC cell
CC proliferation.
XX
SQ Sequence 16 AA;

```
  Query Match             52.2%;  Score 84;  DB 21;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-05;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      13 NNATFYFKIDNVKKAR 28
```

```
          ||||||||||||||||
Db      1 nnatfyfkidnvkkar 16


RESULT   4
AAY95409
ID   AAY95409 standard; Peptide; 16 AA.
XX
AC   AAY95409;
XX
DT   25-SEP-2000  (first entry)
XX
DE   Anti-angiogenic D3 peptide.
XX
KW   Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;
KW   endothelial cell proliferation; apoptosis; cancer; ocular disorder;
KW   rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic;
KW   therapy; human; D3 peptide.
XX
OS   Homo sapiens.
XX
PN   WO200035407-A2.
XX
PD   22-JUN-2000.
XX
PF   02-DEC-1999;   99WO-US28465.
XX
PR   16-DEC-1998;   98US-0112427.
XX
PA   (UTEM ) UNIV TEMPLE.
PA   (MCCR/) MCCRAE R K.
XX
PI   McCrae RK;
XX
DR   WPI; 2000-442247/38.
XX
PT   Composition for inhibiting angiogenesis and endothelial cell
PT   proliferation, inducing endothelial cell apoptosis and treating cancer,
PT   rheumatoid arthritis, and ocular disorders comprises a kininogen domain
PT   3 analog  -
XX
PS   Claim 6; Page 26; 44pp; English.
XX
CC   The present sequence is that of a D3 peptide derived from human
CC   high mol.wt. kininogen (HK) domain 3 (see AAY95426).  The D3 peptide
CC   inhibits endothelial cell proliferation and thus possesses
CC   anti-angiogenic activity.  It is an example of D3 peptides of the
CC   invention (see AAY95405-26) that are analogues of certain sites in
CC   the HK domain 3, in this case amino acid residues Asn275-Lys282.
CC   The peptides inhibit endothelial cell proliferation and may also
CC   induce endothelial cell apoptosis.  Compositions including the
CC   peptides are used in claimed methods for inhibiting angiogenesis,
CC   inhibiting endothelial cell proliferation, and inducing endothelial
CC   cell apoptosis.  Cancer, rheumatoid arthritis, and ocular disorders
CC   characterized by undesired vascularization of the retina are treated.
CC   The IC50 value for the present peptide was less than 0.8 uM for
CC   inhibition of fibroblast growth factor-induced HUVEC cell
CC   proliferation.
XX
SQ   Sequence   16 AA;


  Query Match             51.6%;  Score 83;  DB 21;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-05;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 TITKLNAENNATFYFK 20
          ||||||||||||||||
Db      1 titklnaennatfyfk 16


RESULT   5
AAB08553
ID   AAB08553 standard; Peptide; 16 AA.
XX
AC   AAB08553;
XX
DT   20-DEC-2000  (first entry)
XX
DE   Peptide identified from an origin of prepro-bradykinine.
XX
KW   Precursor peptide; polypeptide hormone; peptide identification.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "hydrogen attached"
FT   Modified-site   16
FT                   /note= "amidated residue"
XX
PN   WO200050636-A1.
XX
PD   31-AUG-2000.
XX
PF   24-FEB-2000; 2000WO-FR00460.
XX
PR   25-FEB-1999;   99US-0257525.
XX
PA   (SCRC ) SCRAS SOC CONSEILS RECH & APPL SCI.
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Camara Ferrer YJA,  Thurieau C,  Martinez J,  Berge G,  Goze C;
XX
DR   WPI; 2000-572101/53.
XX
PT   Identifying peptide with selected function, useful particularly for
PT   C-amidated hormones, by screening database for combination of nucleic
PT   acid and amino acid sequences  -
XX
PS   Claim 16; Page 20; 40pp; French.
XX
CC   The specification describes a method for identifying a peptide having
CC   a particular function. The method comprises preparing a database of
CC   polynucleotides and polypeptides of unknown functions, screening the
CC   database for a combination of nucleotides or amino acids indicative of
CC   the peptide with a particular function, and identifying polynucleotides
CC   and proteins which contain the peptide. The method is used to identify
CC   precursor peptides with an amidated C-terminus, especially polypeptide
CC   hormones, for studying physiologically active substances. The present
CC   sequence represents a peptide which was identified using the method of
CC   the invention.
XX
SQ   Sequence   16 AA;


  Query Match             46.0%;  Score 74;  DB 21;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.0012;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     18 YFKIDNVKKARVQVV 32
          |||||||||||||||
Db      1 yfkidnvkkarvqvv 15


RESULT   6
AAY95406
ID   AAY95406 standard; Peptide; 12 AA.
XX
AC   AAY95406;
XX
DT   25-SEP-2000  (first entry)
XX
DE   Anti-angiogenic peptide N-terminal fragment.
XX
KW   Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;
```

KW endothelial cell proliferation; apoptosis; cancer; ocular disorder;
KW rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic;
KW therapy; human; D3 peptide.
XX
OS Homo sapiens.
XX
PN WO200035407-A2.
XX
PD 22-JUN-2000.
XX
PF 02-DEC-1999; 99WO-US28465.
XX
PR 16-DEC-1998; 98US-0112427.
XX
PA (UTEM) UNIV TEMPLE.
PA (MCCR/) MCCRAE R K.
XX
PI McCrae RK;
XX
DR WPI; 2000-442247/38.
XX
PT Composition for inhibiting angiogenesis and endothelial cell
PT proliferation, inducing endothelial cell apoptosis and treating cancer,
PT rheumatoid arthritis, and ocular disorders comprises a kininogen domain
PT 3 analog -
XX
PS Claim 3; Page 25; 44pp; English.
XX
CC The present sequence is that of an N-terminal fragment of a novel
CC anti-angiogenic D3 peptide (see AAY95408) derived from human high
CC mol.wt. kininogen (HK) domain 3 (see AAY95426). The full-length D3
CC peptide inhibits endothelial cell proliferation and thus possesses
CC anti-angiogenic activity. It is an example of peptides of the
CC invention (see AAY95405-26) that are analogues of certain sites in
CC the HK domain 3. The peptides inhibit endothelial cell proliferation
CC and may also induce endothelial cell apoptosis. Compositions
CC including the peptides are used in claimed methods for inhibiting
CC angiogenesis, inhibiting endothelial cell proliferation, and
CC inducing endothelial cell apoptosis. Cancer, rheumatoid arthritis,
CC and ocular disorders characterized by undesired vascularization of
CC the retina are treated.
XX
SQ Sequence 12 AA;

Query Match 37.3%; Score 60; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.072;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLTHTITKLNAE 12
||||||||||||
Db 1 tlthtitklnae 12

RESULT 7
AAY95407
ID AAY95407 standard; Peptide; 12 AA.
XX
AC AAY95407;
XX
DT 25-SEP-2000 (first entry)
XX
DE Anti-angiogenic peptide C-terminal fragment.
XX
KW Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;
KW endothelial cell proliferation; apoptosis; cancer; ocular disorder;
KW rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic;
KW therapy; human; D3 peptide.
XX
OS Homo sapiens.
XX
PN WO200035407-A2.
XX
PD 22-JUN-2000.
XX
PF 02-DEC-1999; 99WO-US28465.
XX
PR 16-DEC-1998; 98US-0112427.
XX
PA (UTEM) UNIV TEMPLE.
PA (MCCR/) MCCRAE R K.
XX
PI McCrae RK;
XX
DR WPI; 2000-442247/38.
XX
PT Composition for inhibiting angiogenesis and endothelial cell
PT proliferation, inducing endothelial cell apoptosis and treating cancer,
PT rheumatoid arthritis, and ocular disorders comprises a kininogen domain
PT 3 analog -
XX
PS Claim 3; Page 25; 44pp; English.
XX
CC The present sequence is that of a C-terminal fragment of a novel
CC anti-angiogenic D3 peptide (see AAY95408) derived from human high
CC mol.wt. kininogen (HK) domain 3 (see AAY95426). The full-length D3
CC peptide inhibits endothelial cell proliferation and thus possesses
CC anti-angiogenic activity. It is an example of peptides of the
CC invention (see AAY95405-26) that are analogues of certain sites in
CC the HK domain 3. The peptides inhibit endothelial cell proliferation
CC and may also induce endothelial cell apoptosis. Compositions
CC including the peptides are used in claimed methods for inhibiting
CC angiogenesis, inhibiting endothelial cell proliferation, and
CC inducing endothelial cell apoptosis. Cancer, rheumatoid arthritis,
CC and ocular disorders characterized by undesired vascularization of
CC the retina are treated.
XX
SQ Sequence 12 AA;

Query Match 34.8%; Score 56; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 IDNVKKARVQVV 32
||||||||||||
Db 1 idnvkkarvqvv 12

RESULT 8
AAB37455
ID AAB37455 standard; peptide; 9 AA.
XX
AC AAB37455;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human kininogen D3 peptide fragment.
XX
KW Enzyme; legumain; endopeptidase; cystatin; human; kininogen.
XX
OS Homo sapiens.
XX
PN WO200064945-A1.
XX
PD 02-NOV-2000.
XX
PF 20-APR-2000; 2000WO-GB01571.
XX
PR 22-APR-1999; 99GB-0009133.
XX
PA (BABR-) BABRAHAM INST.
XX
PI Abrahamson M, Barrett AJ;
XX
DR WPI; 2000-687316/67.

```
XX
PT   Inhibition of mammalian legumain or legumain-related endopeptidase by
PT   cystatin involves interaction with second papain-non-reactive site of
PT   cystatin -
XX
PS   Disclosure; Fig 4; 45pp; English.
XX
CC   The present invention relates to inhibition of the enzymatic activity of
CC   legumain or a legumain-related endopeptidase by cystatin. The inhibition
CC   involves an interaction between legumain and a papain-non-reactive site
CC   of cystatin. Legumain (EC 3.4.22.34) is a cysteine endopeptidase, and
CC   performs a protein-processing function. The present sequence is a peptide
CC   fragment of human kininogen D3, which was used in the present invention.
CC   Kininogen is a type 3 cystatin. The present sequence is thought to be
CC   involved in a legumain-inhibitory site.
XX
SQ   Sequence   9 AA;


  Query Match             30.4%;  Score 49;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 6.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      13 NNATFYFKI 21
           |||||||||
Db       1 nnatfyfki 9


RESULT   9
AAY95405
ID   AAY95405 standard; Peptide; 10 AA.
XX
AC   AAY95405;
XX
DT   25-SEP-2000  (first entry)
XX
DE   Anti-angiogenic D3 peptide.
XX
KW   Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;
KW   endothelial cell proliferation; apoptosis; cancer; ocular disorder;
KW   rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic;
KW   therapy; human; D3 peptide.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "Xaa represents 0 amino acids, or the
FT                          sequence of AAY95406 or its N-terminal
FT                          truncation fragment containing at least
FT                          1 amino acid"
FT   Peptide         2..9
FT                   /note= "corresponds to residues Asn(275)-Lys(282)
FT                          of HK domain 3"
FT   Misc-difference 10
FT                   /note= "Xaa represents 0 amino acids, or the
FT                          sequence of AAY95407 or its C-terminal
FT                          truncation fragment containing at least
FT                          1 amino acid"
XX
PN   WO200035407-A2.
XX
PD   22-JUN-2000.
XX
PF   02-DEC-1999;   99WO-US28465.
XX
PR   16-DEC-1998;   98US-0112427.
XX
PA   (UTEM ) UNIV TEMPLE.
PA   (MCCR/) MCCRAE R K.
XX
PI   McCrae RK;
XX
DR   WPI; 2000-442247/38.
XX
PT   Composition for inhibiting angiogenesis and endothelial cell
PT   proliferation, inducing endothelial cell apoptosis and treating cancer,
PT   rheumatoid arthritis, and ocular disorders comprises a kininogen domain
PT   3 analog -
XX
PS   Claim 1; Page 25; 44pp; English.
XX
CC   The present sequence is that of a D3 peptide derived from high
CC   mol.wt. kininogen (HK) domain 3 (see AAY95426). The D3 peptide, which
CC   may optionally include N-terminal and/or C-terminal protecting
CC   groups, inhibits endothelial cell proliferation and thus possesses
CC   anti-angiogenic activity. It is an example of peptides of the
CC   invention (see AAY95405-26) which are analogues of certain sites in
CC   the HK domain 3, in this case amino acids Asn275-Lys282. The
CC   peptides inhibit endothelial cell proliferation and may also induce
CC   endothelial cell apoptosis. Compositions including such peptides
CC   are used in claimed methods for inhibiting angiogenesis, inhibiting
CC   endothelial cell proliferation, and inducing endothelial cell
CC   apoptosis. Cancer, rheumatoid arthritis, and ocular disorders
CC   characterized by undesired vascularization of the retina are treated.
XX
SQ   Sequence   10 AA;


  Query Match             28.0%;  Score 45;  DB 21;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 6.5;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      13 NNATFYFK 20
           ||||||||
Db       2 nnatfyfk 9


RESULT  10
AAW94743
ID   AAW94743 standard; Protein; 32 AA.
XX
AC   AAW94743;
XX
DT   22-APR-1999  (first entry)
XX
DE   Anti-Staph (HAY) 96-110 light chain variable region.
XX
KW   Monoclonal antibody; MAb; lipoteichoic acid; gram positive; bacteria;
KW   immunoglobulin; phagocytosis; infection; epitope; peptide mimic;
KW   MAb 96-110.
XX
OS   Mus sp.
XX
PN   WO9857994-A2.
XX
PD   23-DEC-1998.
XX
PF   16-JUN-1998;   98WO-US12402.
XX
PR   16-JUN-1997;   97US-0049871.
XX
PA   (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.
XX
PI   Fischer GW,  Schuman RF,  Stinson JL,  Wong H;
XX
DR   WPI; 1999-095329/08.
DR   N-PSDB; AAX05587.
XX
PT   New antibodies to lipoteichoic acid of gram positive bacteria - used
PT   to develop products for the diagnosis, prevention and treatment of
PT   infections caused by gram positive bacteria
XX
PS   Claim 21; Fig 12; 150pp; English.
XX
CC   The invention relates to a monoclonal antibody (MAb) to lipoteichoic
```

```
CC   acid of gram positive bacteria, where the MAb is a chimeric
CC   immunoglobulin comprising at least part of a human immunoglobulin
CC   constant region and at least part of a non-human immunoglobulin variable
CC   region having specificity to lipoteichoic acid of gram positive bacteria.
CC   The antibodies bind to whole bacteria and enhance phagocytosis and
CC   killing of the bacteria and enhance protection from lethal infection. The
CC   antibodies or peptides (encoded by a DNA of the variable region of
CC   anti-lipoteichoic acid antibody or characterised by amino acids
CC   corresponding to one or more of the complementarity determining regions
CC   (CDRs) of the variable region of the antibody) can be used for treating
CC   or preventing infections caused by gram positive bacteria. They can also
CC   be used for the diagnosis of gram positive bacterial infections.
CC   Sequences AAW94740-44 represent light chain variable regions of the
CC   anti-lipoteichoic antibody 96-100.
XX
SQ   Sequence   32 AA;


  Query Match             24.8%;  Score 40;  DB 20;  Length 32;
  Best Local Similarity   50.0%;  Pred. No. 1.2e+02;
  Matches    7;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 TITKLNAENNATFY 18
           ||::: ||: ||:|
Db      18 tisrveaedaatyy 31


RESULT  11
AAR27043
ID   AAR27043 standard; peptide; 32 AA.
XX
AC   AAR27043;
XX
DT   01-MAR-1993  (first entry)
XX
DE   Human light chain framework 3 used in humanised MAb.
XX
KW   Monoclonal antibody; complementarity determining region; CDR;
KW   antigens; tumour; melanoma; carcinoma; glioma.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 15
FT                   /label= Tyr, Phe, Trp, His
XX
PN   WO9215683-A.
XX
PD   17-SEP-1992.
XX
PF   04-MAR-1992;   92WO-EP00480.
XX
PR   06-MAR-1991;   91EP-0103389.
XX
PA   (MERE ) MERCK PATENT GMBH.
XX
PI   Bendig MM,  Kettleborough CA,  Saldanha J;
XX
DR   WPI; 1992-331729/40.
XX
PT   Human monoclonal antibodies binding to human receptors - for
PT   treatment and diagnosis of tumours, e.g. melanoma and carcinoma
XX
PS   Claim 5; Page 63; 89pp; English.
XX
CC   The human framework sequence FR3 was used in the prodn. of
CC   humanised monoclonal antibodies comprising antigen binding sites
CC   (CDRs) of non human origen based on consensus sequence of CDRs.
CC   The sequence is a light chain framework region. The humanised
CC   antibodies or their chimeric variants may be used as therapeutic
CC   or diagnostic agents in order to combat e.g. glioma, melanoma or
CC   carcinoma.    See also AAR27299-300 and AAR27037-51.
XX
SQ   Sequence   32 AA;


  Query Match             24.2%;  Score 39;  DB 13;  Length 32;
  Best Local Similarity   50.0%;  Pred. No. 1.6e+02;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 THTITKLNAENNATFY 18
           | ||: |  |: ||:|
Db      16 tftisslqpediatyy 31


RESULT  12
AAR85155
ID   AAR85155 standard; Protein; 32 AA.
XX
AC   AAR85155;
XX
DT   18-JAN-1996  (first entry)
XX
DE   Human ONS-M21 antibody light variable region FR 3.
XX
KW   Human; ONS-M21 antibody; light variable region; chimeric protein;
KW   framework region; FR 3; medulloblastoma; brain tumour; treatment;
KW   diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO9514041-A1.
XX
PD   26-MAY-1995.
XX
PF   19-OCT-1994;   94WO-JP01763.
XX
PR   19-NOV-1993;   93JP-0291078.
XX
PA   (CHUS ) CHUGAI SEIYAKU KK.
XX
PI   Ohtomo T,  Sato K,  Tsuchiya M;
XX
DR   WPI; 1995-200347/26.
XX
PT   Reconstituted antibody against human medullo:blastoma cells -
PT   contains high proportion of human antibody origin and has low
PT   antigenicity
XX
PS   Claim 9; Page 101; 120pp; Japanese.
XX
CC   AAR85153-R85156 are human antibody ONS-M21 light variable region
CC   framework regions (FRs). They were used in the construction of
CC   a human/murine chimeric antibody, reactive with human medullo-
CC   blastoma (a brain tumour) cells. The chimeric antibody can be
CC   used in the diagnosis and treatment of this disease.
XX
SQ   Sequence   32 AA;


  Query Match             24.2%;  Score 39;  DB 16;  Length 32;
  Best Local Similarity   50.0%;  Pred. No. 1.6e+02;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 THTITKLNAENNATFY 18
           | ||: |  |: ||:|
Db      16 tftisslqpediatyy 31


RESULT  13
AAR77206
ID   AAR77206 standard; Peptide; 32 AA.
XX
AC   AAR77206;
XX
DT   23-AUG-1995  (first entry)
```

XX
DE Mouse anti-human IL-6 Ab L chain V region framework region 3.
XX
KW Primer; PCR; amplify; kappa; light chain; variable region; mouse; human;
KW interleukin; antibody; hybridoma; CDR; framework; constant region;
KW heavy chain; disorder; antigenicity.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 15
FT /label= Phe or Tyr
FT /*note= "In HEF-RVL-SK2a, this position is Phe;
FT in HEF-RVL-SK2b, this position is Tyr"
XX
PN WO9428159-A.
XX
PD 08-DEC-1994.
XX
PF 30-MAY-1994; 94WO-JP00859.
XX
PR 31-MAY-1993; 93JP-0129787.
XX
PA (CHUS) CHUGAI SEIYAKU KK.
PA (CHUS) CHUGAI PHARM CO LTD.
XX
PI Hirata Y, Sato K, Tsuchiya M;
XX
DR WPI; 1995-022828/03.
XX
PT Antibody against IL-6 - useful for the therapy and treatment of
PT IL-6 related disorders.
XX
PS Claim 6; Page 64; 82pp; Japanese.
XX
CC The sequence of the mouse anti-human interleukin-6 (IL-6) antibody light
CC chain variable region framework region (FR) 3. The sequences of FR1-4
CC (AAR77204-7) were used in conjunction with the complementarity
CC determining regions 1-3 (AAR77201-3) to construct a chimaeric antibody
CC against human interleukin-6 (IL-6). The vectors AAQ75914-7 express
CC constructs encoding fragments of a chimaeric antibody to the human IL-6
CC comprising (a) a light chain with (i) a variable region containing 3 CDR
CC (AAR77201-3) inserted into several framework regions (FR) (AAR77204-7)
CC and (ii) a human light chain constant region and (b) a heavy chain with
CC (i) a variable region containing 3 CDR (AAR77212-4) inserted into FR
CC (AAR77215-8) and (ii) a human light chain constant region. The FR of the
CC light chain may be mouse derived (AAQ75888) or from the human antibody
CC REI. The heavy chain FR may also be mouse derived (AAQ75889) or from the
CC human antibody DAW. The antibodies can be used in the treatment of IL-6
CC related disorders. The antibodies are useful as they have low
CC antigenicity due to the use of human derived sequences and low
CC antigenicity mouse derived sequences.
XX
SQ Sequence 32 AA;

Query Match 24.2%; Score 39; DB 16; Length 32;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 3 THTITKLNAENNATFY 18
| ||: | |: ||:|
Db 16 tftisslqpediatyy 31

RESULT 14
AAR96288
ID AAR96288 standard; peptide; 32 AA.
XX
AC AAR96288;
XX
DT 07-JAN-1997 (first entry)
XX
DE Light chain framework region 3 for human EGF-R antibody.
XX
KW Antibody; Ab; human; EGF-R; epidermal growth factor receptor; monoclonal;
KW humanised antibody; framework region; antigen binding site; light chain;
KW CDR; heavy chain; gamma-1 chain; kappa chain; immunoglobulin; epithelium;
KW malignant tumour; breast; bladder; ovary; colon; lung; brain; oesophagus;
KW therapy.
XX
OS Synthetic.
XX
PN EP712863-A1.
XX
PD 22-MAY-1996.
XX
PF 15-NOV-1995; 95EP-0203126.
XX
PR 18-NOV-1994; 94CU-0000128.
XX
PA (IMMU-) CENT IMMUNOLOGIA MOLECULAR.
XX
PI Frias EM, Mateo Del Acosta Del Rio CM, Rodriquez RP;
XX
DR WPI; 1996-252836/26.
XX
PT Humanised and chimeric antibodies for EGF receptor - used in
PT diagnosis and therapy of tumours
XX
PS Claim 6; Page 12; 24pp; English.
XX
CC AAR96286-R96293 represent framework regions (FR) for a humanised
CC monoclonal antibody (Ab) specific for the human epidermal growth factor
CC (EGF) receptor. This Ab binds to human EGF-R, and inhibits binding of
CC EGF to the EGF-R. The humanised Ab comprises antigen binding sites
CC (CDR's) of non-human origin, and FRs of the variable and constant regions
CC of light and heavy chains of human origin. The constant heavy chain
CC region used comprises the amino acid sequence of a gamma-1 chain. The
CC constant light chain region used comprises the kappa chain of a human
CC immunoglobulin. High levels of EGF-R have been detected in malignant
CC tumours originating in the epithelium. These include breast tumours,
CC bladder tumours, ovarian tumours, colonic tumours, lung tumours, brain
CC tumours and tumours of the oesophagus. The presence of EGF-R in tumour
CC cells is an indicator of a poor prognosis (particularly in human breast
CC cancer). The Ab can therefore be used for diagnostic localisation and
CC assessment of tumour growth. The Ab can also be used in the design of a
CC drug targetted to the tumour.
XX
SQ Sequence 32 AA;

Query Match 24.2%; Score 39; DB 17; Length 32;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 3 THTITKLNAENNATFY 18
| ||: | |: ||:|
Db 16 tftisslqpediatyy 31

RESULT 15
AAR97321
ID AAR97321 standard; peptide; 32 AA.
XX
AC AAR97321;
XX
DT 15-OCT-1996 (first entry)
XX
DE Humanised monoclonal antibody light chain framework region.
XX
KW Monoclonal antibody; humanised; mouse; framework region; FR; CDR;
KW complementary determining region; anti-carcinoembryonic antigen;
KW CEA; diagnosis; imaging; therapy; immune response.
XX
OS Homo sapiens.

```
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 4
FT                   /note= "Serine or Aspartic acid."
FT   Misc-difference 10
FT                   /note= "Glycine or Valine."
XX
PN   WO9611013-A1.
XX
PD   18-APR-1996.
XX
PF   28-SEP-1995;   95WO-US11964.
XX
PR   05-OCT-1994;   94US-0318157.
XX
PA   (IMMU-) IMMUNOMEDICS INC.
XX
PI   Armour KL,  Hansen HJ;
XX
DR   WPI; 1996-209653/21.
XX
PT   New humanised anti-CEA monoclonal antibody - having engrafted murine
PT   CDRs, used for diagnosis, imaging and therapy of CEA-producing
PT   cancers
XX
PS   Claim 7; Page 39; 62pp; English.
XX
CC   New humanised monoclonal antibodies (MAbs) comprising the
CC   complementary determining regions (CDRs) of a parental murine class
CC   III anti-carcinoembryonic (CEA) MAb engrafted to the framework
CC   regions (FRs) of a heterologous antibody which can be derived from
CC   any species including human, retain the anti-CEA binding specificity
CC   of the parental murine MAb but are less immunogenic in a human
CC   subject than the parental MAb.  The humanised antibodies can be used
CC   in diagnosis, imaging and therapy of CEA-producing cancers and
CC   patients receiving the humanised antibodies and conjugates show
CC   improved therapeutic results, decreased immune responses and
CC   decreased immune-mediated adverse effects compared to the parent
CC   antibody.  This sequence corresponds to the third framework region
CC   of the light chain of the humanised MAb.  See AAR97313-97333.
XX
SQ   Sequence   32 AA;


  Query Match             24.2%;  Score 39;  DB 17;  Length 32;
  Best Local Similarity   50.0%;  Pred. No. 1.6e+02;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 THTITKLNAENNATFY 18
           | ||: |  |: ||:|
Db      16 tftisslqpediatyy 31


RESULT  16
AAR87043
ID   AAR87043 standard; Peptide; 32 AA.
XX
AC   AAR87043;
XX
DT   25-JUN-1996  (first entry)
XX
DE   Human group I light chain framework 3.
XX
KW   Humanised antibody; interleukin-5; IL-5; recombinant antibody;
KW   antibody engineering; monoclonal antibody; MAb; 39D10; CDR;
KW   complementarity determining region; light chain; framework;
KW   eosinophilia; allergy; asthma.
XX
OS   Homo sapiens.
XX
PN   WO9535375-A1.
XX
PD   28-DEC-1995.
XX
PF   16-JUN-1995;   95WO-GB01411.
XX
PR   17-JUN-1994;   94GB-0012230.
XX
PA   (CLLT ) CELLTECH THERAPEUTICS LTD.
XX
PI   Athwal DS,  Bodmer MW,  Emtage JS;
XX
DR   WPI; 1996-058412/06.
XX
PT   Anti-human IL-5 recombinant antibody - useful for preventing or
PT   reducing eosinophilia and for treating certain allergic diseases,
PT   esp. asthma
XX
PS   Example 3; Fig 3; 69pp; English.
XX
CC   Framework regions (AAR87041-44) of human group I (gp1) germ line
CC   antibody light chain showed homology to corresponding regions
CC   (AAR87045-48, respectively) of the rat anti-human interleukin-5
CC   monoclonal antibody 39D10 light chain (see AAR87040).  This homology
CC   was utilised in the prodn. of a humanised 39D10 VL (AAR87057) in
CC   which rat 39D10 VL complementarity determining regions were grafted
CC   into the human gp1 framework.
XX
SQ   Sequence   32 AA;


  Query Match             23.0%;  Score 37;  DB 17;  Length 32;
  Best Local Similarity   50.0%;  Pred. No. 3.1e+02;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 THTITKLNAENNATFY 18
           | ||: |  |: ||:|
Db      16 tltisslqpedfatyy 31


RESULT  17
AAY52745
ID   AAY52745 standard; Peptide; 32 AA.
XX
AC   AAY52745;
XX
DT   26-JAN-2000  (first entry)
XX
DE   Humanised ATR-5 L chain V region FR3 for "a".
XX
KW   Human tissue factor; TF; humanised; antibody; mouse monoclonal antibody;
KW   ATR-2; ATR-3; ATR-4; ATR-5; ATR-7; ATR-8; thrombotic disease; DIC;
KW   disseminated intravascular coagulation; immunogenicity; chimeric.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9951743-A1.
XX
PD   14-OCT-1999.
XX
PF   02-APR-1999;   99WO-JP01768.
XX
PR   03-APR-1998;   98JP-0091850.
XX
PA   (CHUS ) CHUGAI SEIYAKU KK.
XX
PI   Sato K,  Adachi H,  Yabuta N;
XX
DR   WPI; 1999-620204/53.
XX
PT   Humanised antibody recognizing human tissue factor, used for treatment
PT   of disseminated intravascular coagulation  -
XX
PS   Claim 17; Page 270; 291pp; Japanese.
XX
```

CC The present invention describes chimeric antibody (Ab) heavy (H) chains
CC containing the variable region of the H chain of a mouse monoclonal Ab
CC recognising human tissue factor (hTF) and the constant region of the H
CC chain of a human Ab. The variable region is one of six specified
CC sequences (which are the H chain variable regions from mouse monoclonal
CC Ab's ATR-2,3,4,5,7 or 8). Also described are chimeric Ab light (L)
CC chains containing the variable region of the L chain of a mouse
CC monoclonal Ab recognising human tissue factor (hTF) and the constant
CC region of the L chain of a human Ab, the variable region being one of six
CC specified sequences (which are the L chain variable regions from mouse
CC monoclonal Ab's ATR-2,3,4,5,7 or 8). The chimeric Ab's can be used for
CC the treatment and prevention of thrombotic disease, especially of
CC disseminated intravascular coagulation (DIC). The humanised antibody has
CC the high hTF binding activity of the mouse monoclonal antibody but
CC greatly reduced immunogenicity. AAZ33001 to AAZ33091 and Y527007 to
CC AAY52767 represent sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 32 AA;

Query Match 23.0%; Score 37; DB 20; Length 32;
Best Local Similarity 50.0%; Pred. No. 3.1e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 3 THTITKLNAENNATFY 18
| ||: | |: ||:|
Db 16 tltisslqpedfatyy 31

RESULT 18
AAY52746
ID AAY52746 standard; Peptide; 32 AA.
XX
AC AAY52746;
XX
DT 26-JAN-2000 (first entry)
XX
DE Humanised ATR-5 L chain V region FR3 for "b", "b1" and "b2".
XX
KW Human tissue factor; TF; humanised; antibody; mouse monoclonal antibody;
KW ATR-2; ATR-3; ATR-4; ATR-5; ATR-7; ATR-8; thrombotic disease; DIC;
KW disseminated intravascular coagulation; immunogenicity; chimeric.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9951743-A1.
XX
PD 14-OCT-1999.
XX
PF 02-APR-1999; 99WO-JP01768.
XX
PR 03-APR-1998; 98JP-0091850.
XX
PA (CHUS) CHUGAI SEIYAKU KK.
XX
PI Sato K, Adachi H, Yabuta N;
XX
DR WPI; 1999-620204/53.
XX
PT Humanised antibody recognizing human tissue factor, used for treatment
PT of disseminated intravascular coagulation -
XX
PS Claim 17; Page 270; 291pp; Japanese.
XX
CC The present invention describes chimeric antibody (Ab) heavy (H) chains
CC containing the variable region of the H chain of a mouse monoclonal Ab
CC recognising human tissue factor (hTF) and the constant region of the H
CC chain of a human Ab. The variable region is one of six specified
CC sequences (which are the H chain variable regions from mouse monoclonal
CC Ab's ATR-2,3,4,5,7 or 8). Also described are chimeric Ab light (L)
CC chains containing the variable region of the L chain of a mouse
CC monoclonal Ab recognising human tissue factor (hTF) and the constant
CC region of the L chain of a human Ab, the variable region being one of six
CC specified sequences (which are the L chain variable regions from mouse
CC monoclonal Ab's ATR-2,3,4,5,7 or 8). The chimeric Ab's can be used for
CC the treatment and prevention of thrombotic disease, especially of
CC disseminated intravascular coagulation (DIC). The humanised antibody has
CC the high hTF binding activity of the mouse monoclonal antibody but
CC greatly reduced immunogenicity. AAZ33001 to AAZ33091 and Y527007 to
CC AAY52767 represent sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 32 AA;

Query Match 23.0%; Score 37; DB 20; Length 32;
Best Local Similarity 50.0%; Pred. No. 3.1e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 3 THTITKLNAENNATFY 18
| ||: | |: ||:|
Db 16 tltisslqpedfatyy 31

RESULT 19
AAY52747
ID AAY52747 standard; Peptide; 32 AA.
XX
AC AAY52747;
XX
DT 26-JAN-2000 (first entry)
XX
DE Humanised ATR-5 L chain V region FR3 for "c".
XX
KW Human tissue factor; TF; humanised; antibody; mouse monoclonal antibody;
KW ATR-2; ATR-3; ATR-4; ATR-5; ATR-7; ATR-8; thrombotic disease; DIC;
KW disseminated intravascular coagulation; immunogenicity; chimeric.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9951743-A1.
XX
PD 14-OCT-1999.
XX
PF 02-APR-1999; 99WO-JP01768.
XX
PR 03-APR-1998; 98JP-0091850.
XX
PA (CHUS) CHUGAI SEIYAKU KK.
XX
PI Sato K, Adachi H, Yabuta N;
XX
DR WPI; 1999-620204/53.
XX
PT Humanised antibody recognizing human tissue factor, used for treatment
PT of disseminated intravascular coagulation -
XX
PS Claim 17; Page 271; 291pp; Japanese.
XX
CC The present invention describes chimeric antibody (Ab) heavy (H) chains
CC containing the variable region of the H chain of a mouse monoclonal Ab
CC recognising human tissue factor (hTF) and the constant region of the H
CC chain of a human Ab. The variable region is one of six specified
CC sequences (which are the H chain variable regions from mouse monoclonal
CC Ab's ATR-2,3,4,5,7 or 8). Also described are chimeric Ab light (L)
CC chains containing the variable region of the L chain of a mouse
CC monoclonal Ab recognising human tissue factor (hTF) and the constant
CC region of the L chain of a human Ab, the variable region being one of six
CC specified sequences (which are the L chain variable regions from mouse
CC monoclonal Ab's ATR-2,3,4,5,7 or 8). The chimeric Ab's can be used for
CC the treatment and prevention of thrombotic disease, especially of
CC disseminated intravascular coagulation (DIC). The humanised antibody has
CC the high hTF binding activity of the mouse monoclonal antibody but

CC greatly reduced immunogenicity. AAZ33001 to AAZ33091 and Y527007 to
CC AAY52767 represent sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 32 AA;

Query Match 23.0%; Score 37; DB 20; Length 32;
Best Local Similarity 50.0%; Pred. No. 3.1e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 3 THTITKLNAENNATFY 18
| ||: | |: ||:|
Db 16 tltisslqpediatyy 31

RESULT 20
AAB98286
ID AAB98286 standard; Peptide; 32 AA.
XX
AC AAB98286;
XX
DT 20-AUG-2001 (first entry)
XX
DE Anti-A33 antigen immunoglobulin VL FR3 SEQ ID NO:92.
XX
KW Human; rabbit; humanised; A33 antigen; anti-A33 antigen antibody;
KW immunoglobulin; complementarity determining region; CDR; cancer;
KW cytostatic; anticancer; colon cancer; stomach cancer.
XX
OS Homo sapiens.
XX
PN WO200130393-A2.
XX
PD 03-MAY-2001.
XX
PF 20-OCT-2000; 2000WO-US29289.
XX
PR 22-OCT-1999; 99US-0425638.
PR 04-APR-2000; 2000US-0543004.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
PA (SLOK) SLOAN KETTERING INST CANCER RES.
PA (SCRI) SCRIPPS RES INST.
XX
PI Barbas CF, Rader C, Ritter G, Welt S, Old LJ;
XX
DR WPI; 2001-328613/34.
XX
PT Treating cancers, particularly of stomach and colon, that express A33
PT antigen by administering conjugate of anticancer agent with specific
PT immunoglobulin product -
XX
PS Claim 16; Page 40; 85pp; English.
XX
CC The present invention describes a method for treating cancers that
CC express the A33 antigen. The method comprises administering an
CC anticancer agent (I) conjugated to an immunoglobulin product (II) that
CC binds specifically to A33 and contains one or more of 13 specified
CC complementarity determining regions (CDRs), given in AAB98262 to
CC AAB98274. (I) has cytostatic activity. The method can be used for
CC treating colon and stomach cancers. (II), or the nucleic acid encoding
CC it, can be used directly, in unconjugated form, for immunotherapy of
CC cancer, and, when labeled, for detection or diagnosis of diseases
CC associated with A33 expression. AAH22218 to AAH22254 and AAB98230 to
CC AAB98321 represent sequences used in the exemplification of the
CC present invention.
XX
SQ Sequence 32 AA;

Query Match 23.0%; Score 37; DB 22; Length 32;
Best Local Similarity 50.0%; Pred. No. 3.1e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 3 THTITKLNAENNATFY 18
| ||: | |: ||:|
Db 16 tltisslqpedvatyy 31

Search completed: July 8, 2002, 11:49:53
Job time: 154 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 8, 2002, 11:37:53 ; Search time 12.96 Seconds
(without alignments)
15.078 Million cell updates/sec

Title: US-09-461-061A-1
Perfect score: 45
Sequence: 1 NNATFYFK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 48605

Minimum DB seq length: 0
Maximum DB seq length: 8

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	22	48.9	6	1	US-08-336-618-8	Sequence 8, Appli
2	22	48.9	7	4	US-09-084-605B-14	Sequence 14, Appl
3	21	46.7	4	4	US-08-669-304-5	Sequence 5, Appli
4	21	46.7	4	4	US-08-669-304-12	Sequence 12, Appl
5	20	44.4	7	1	US-08-302-449-11	Sequence 11, Appl
6	20	44.4	7	5	PCT-US94-07430-11	Sequence 11, Appl
7	20	44.4	8	4	US-08-160-604-7	Sequence 7, Appli
8	19	42.2	4	4	US-09-177-249-143	Sequence 143, App
9	19	42.2	6	2	US-08-282-980B-2	Sequence 2, Appli
10	19	42.2	6	2	US-08-282-980B-3	Sequence 3, Appli
11	19	42.2	6	2	US-08-282-980B-4	Sequence 4, Appli
12	19	42.2	6	2	US-08-482-228-170	Sequence 170, App
13	19	42.2	6	3	US-08-931-095-2	Sequence 2, Appli
14	19	42.2	6	3	US-08-931-095-7	Sequence 7, Appli
15	19	42.2	6	3	US-08-931-095-9	Sequence 9, Appli
16	19	42.2	6	3	US-08-482-528-170	Sequence 170, App
17	19	42.2	7	1	US-08-290-448A-82	Sequence 82, Appl
18	19	42.2	7	1	US-08-290-448A-82	Sequence 82, Appl
19	19	42.2	7	1	US-08-487-890A-81	Sequence 81, Appl
20	19	42.2	7	1	US-08-175-069A-82	Sequence 82, Appl
21	19	42.2	7	2	US-08-347-397-15	Sequence 15, Appl
22	19	42.2	7	2	US-08-871-163-39	Sequence 39, Appl
23	19	42.2	7	2	US-08-478-435-81	Sequence 81, Appl
24	19	42.2	7	2	US-08-337-483-81	Sequence 81, Appl
25	19	42.2	7	2	US-08-478-373-81	Sequence 81, Appl
26	19	42.2	7	2	US-08-482-228-30	Sequence 30, Appl
27	19	42.2	7	3	US-08-474-671-81	Sequence 81, Appl
28	19	42.2	7	3	US-08-483-577A-81	Sequence 81, Appl
29	19	42.2	7	3	US-08-931-095-3	Sequence 3, Appli
30	19	42.2	7	3	US-08-482-528-30	Sequence 30, Appl
31	19	42.2	7	3	US-08-767-903-39	Sequence 39, Appl
32	19	42.2	7	4	US-08-897-438-81	Sequence 81, Appl
33	19	42.2	7	4	US-08-461-939B-82	Sequence 82, Appl
34	19	42.2	7	4	US-08-464-000-82	Sequence 82, Appl
35	19	42.2	7	4	US-08-637-654-81	Sequence 81, Appl
36	19	42.2	8	3	US-08-696-854B-10	Sequence 10, Appl
37	19	42.2	8	5	PCT-US91-02942-85	Sequence 85, Appl
38	18	40.0	5	1	US-07-924-753-6	Sequence 6, Appli
39	18	40.0	5	1	US-08-068-947-2	Sequence 2, Appli
40	18	40.0	5	4	US-09-066-481-23	Sequence 23, Appl
41	18	40.0	5	4	US-09-357-952-133	Sequence 133, App
42	18	40.0	5	4	US-09-521-650-133	Sequence 133, App
43	18	40.0	5	4	US-09-168-888-133	Sequence 133, App
44	18	40.0	6	2	US-08-377-432-37	Sequence 37, Appl
45	18	40.0	6	2	US-08-871-163-38	Sequence 38, Appl

ALIGNMENTS

```
RESULT   1
US-08-336-618-8
; Sequence 8, Application US/08336618
; Patent No. 5763590
;  GENERAL INFORMATION:
;    APPLICANT:  Peattie, Debra A.
;    APPLICANT:  Harding, Matthew W.
;    APPLICANT:  Livingston, David J.
;    TITLE OF INVENTION:  ISOLATION OF AN Mr 52,000 FK506 BINDING
;    TITLE OF INVENTION:  PROTEIN AND MOLECULAR CLONING OF A CORRESPONDING HUMAN
;    TITLE OF INVENTION:  cDNA
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith and Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/336,618
;      FILING DATE:  09-NOV-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/963,325
;      FILING DATE:  16-OCT-1992
;      APPLICATION NUMBER:  US 07/777,752
;      FILING DATE:  11-OCT-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/
;      FILING DATE:  09-OCT-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  VPI91-06A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-336-618-8
```

```
  Query Match             48.9%;  Score 22;  DB 1;  Length 6;
  Best Local Similarity   80.0%;  Pred. No. 1.7e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 TFYFK 8
           | |||
Db       2 TVYFK 6


RESULT   2
US-09-084-605B-14
; Sequence 14, Application US/09084605B
; Patent No. 6329501
; GENERAL INFORMATION:
;  APPLICANT:  Smith, Bruce F.
;  APPLICANT:  Samoilova, Tatiana
;  TITLE OF INVENTION:  Methods and Compositions for Targeting
;  TITLE OF INVENTION:  Compounds to Muscle
;  FILE REFERENCE:  5721-8
;  CURRENT APPLICATION NUMBER: US/09/084,605B
;  CURRENT FILING DATE:  1998-05-26
;  NUMBER OF SEQ ID NOS:  29
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 14
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Phage display library peptides
US-09-084-605B-14


  Query Match             48.9%;  Score 22;  DB 4;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+05;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 ATFY 6
           ||||
Db       4 ATFY 7


RESULT   3
US-08-669-304-5
; Sequence 5, Application US/08669304
; Patent No. 6251626
;  GENERAL INFORMATION:
;    APPLICANT:  Peter Stougaard
;    APPLICANT:  Ole Cai Hansen
;    TITLE OF INVENTION:  RECOMBINANT HEXOSE OXIDASE, A
;    TITLE OF INVENTION:  METHOD OF PRODUCING SAME AND USE OF SUCH ENZYME
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hunton & Williams
;      STREET:  1900 K Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20006-1109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/669,304
;      FILING DATE:  12 July 1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/476,910
;      FILING DATE:  7 June 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stanislaus Aksman
;      REGISTRATION NUMBER:  28,562
;      REFERENCE/DOCKET NUMBER:  320.000003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 955-1926
;      TELEFAX:  (202) 778-2201
;      TELEX:  No. 6251626e
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-669-304-5


  Query Match             46.7%;  Score 21;  DB 4;  Length 4;
  Best Local Similarity   75.0%;  Pred. No. 1.7e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 FYFK 8
           :|||
Db       1 YYFK 4


RESULT   4
US-08-669-304-12
; Sequence 12, Application US/08669304
; Patent No. 6251626
;  GENERAL INFORMATION:
;    APPLICANT:  Peter Stougaard
;    APPLICANT:  Ole Cai Hansen
;    TITLE OF INVENTION:  RECOMBINANT HEXOSE OXIDASE, A
;    TITLE OF INVENTION:  METHOD OF PRODUCING SAME AND USE OF SUCH ENZYME
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hunton & Williams
;      STREET:  1900 K Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20006-1109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/669,304
;      FILING DATE:  12 July 1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/476,910
;      FILING DATE:  7 June 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stanislaus Aksman
;      REGISTRATION NUMBER:  28,562
;      REFERENCE/DOCKET NUMBER:  320.000003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 955-1926
;      TELEFAX:  (202) 778-2201
;      TELEX:  No. 6251626e
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-669-304-12
```

Query Match 46.7%; Score 21; DB 4; Length 4;
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy      5 FYFK 8
          :|||
Db      1 YYFK 4
```

RESULT 5
US-08-302-449-11
; Sequence 11, Application US/08302449
; Patent No. 5679635
; GENERAL INFORMATION:
; APPLICANT: Matalon, Reuben
; APPLICANT: Kaul, Rajinder
; APPLICANT: Cao, Guang Ping
; APPLICANT: Balamurugan, Kuppareddi
; APPLICANT: Michals-Matalon, Kimberlee
; TITLE OF INVENTION: Aspartoacylase Gene, Protein, and
; TITLE OF INVENTION: Methods of Screening for Mutations Associated with Canavan
; TITLE OF INVENTION: Disease
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Millen, White, Zelano & Branigan, P.C.
; STREET: 2200 Clarendon Boulevard, Suite 1400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/302,449
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/128,020
; FILING DATE: 29-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/07430
; FILING DATE: 05-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hamlet-King, Diana
; REGISTRATION NUMBER: 33,302
; REFERENCE/DOCKET NUMBER: SHUTT 2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-243-6333
; TELEFAX: 703-243-6410
; TELEX: 64191
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
US-08-302-449-11

Query Match 44.4%; Score 20; DB 1; Length 7;
Best Local Similarity 50.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 NNATFY 6
          | | :|
Db      2 NEAAYY 7
```

RESULT 6
PCT-US94-07430-11
; Sequence 11, Application PC/TUS9407430
; GENERAL INFORMATION:
; APPLICANT: Matalon, Reuben
; APPLICANT: Kaul, Rajinder
; APPLICANT: Cao, Guang Ping
; APPLICANT: Balamurugan, Kuppareddi
; APPLICANT: Michals-Matalon, Kimberlee
; TITLE OF INVENTION: Aspartoacylase Gene, Protein, and
; TITLE OF INVENTION: Methods of Screening for Mutations Associated with Canavan
; TITLE OF INVENTION: Disease
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Millen, White, Zelano & Branigan, P.C.
; STREET: 2200 Clarendon Boulevard, Suite 1400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/07430
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/128,020
; FILING DATE: 29-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hamlet-King, Diana
; REGISTRATION NUMBER: 33,302
; REFERENCE/DOCKET NUMBER: SHUTT 1PO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-243-6333
; TELEFAX: 703-243-6410
; TELEX: 64191
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
PCT-US94-07430-11

Query Match 44.4%; Score 20; DB 5; Length 7;
Best Local Similarity 50.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 NNATFY 6
          | | :|
Db      2 NEAAYY 7
```

RESULT 7
US-08-160-604-7
; Sequence 7, Application US/08160604
; Patent No. 6232522
; GENERAL INFORMATION:
; APPLICANT: Harley, John
; APPLICANT: James, Judith A.
; APPLICANT: Scofield, R. H.
; TITLE OF INVENTION: PEPTIDE INDUCTION OF AUTOIMMUNITYAND CLINICAL SYMPTOMATOLOG
; NUMBER OF SEQUENCES: 127
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia

```
;     COUNTRY:  USA
;     ZIP:  30309-4530
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/160,604
;     FILING DATE:  30-NOV-1993
;     CLASSIFICATION:  424
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 07/867,819
;     FILING DATE:  13-APR-1992
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 07/648,205
;     FILING DATE:  31-JAN-1991
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 07/472,947
;     FILING DATE:  31-JAN-1990
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Pabst, Patrea L.
;     REGISTRATION NUMBER:  31,284
;     REFERENCE/DOCKET NUMBER:  OMRF114CIP(3)
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  (404)-815-6508
;     TELEFAX:  (404)-815-6555
;  INFORMATION FOR SEQ ID NO:  7:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  8 amino acids
;     TYPE:  amino acid
;     STRANDEDNESS:  single
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  peptide
;   HYPOTHETICAL:  NO
;   ANTI-SENSE:  NO
;   FRAGMENT TYPE:  N-terminal
US-08-160-604-7


  Query Match             44.4%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   60.0%;  Pred. No. 1.7e+05;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 TFYFK 8
          |:| |
Db      1 TYYIK 5


RESULT   8
US-09-177-249-143
; Sequence 143, Application US/09177249
; Patent No. 6229064
; GENERAL INFORMATION:
;  APPLICANT: Fischer, Robert L.
;  APPLICANT: Ohad, Nir
;  APPLICANT: Kiyosue, Tomohiro
;  APPLICANT: Yadegari, Ramin
;  APPLICANT: Margossian, Linda
;  APPLICANT: Harada, John
;  APPLICANT: Goldberg, Robert B.
;  APPLICANT: The Regents of the University of California
;  TITLE OF INVENTION: Nucleic Acids That Control Seed and Fruit
;  TITLE OF INVENTION: Development in Plants
;  FILE REFERENCE: 023070-086120US
;  CURRENT APPLICATION NUMBER: US/09/177,249
;  CURRENT FILING DATE: 1998-10-22
;  EARLIER APPLICATION NUMBER: US 09/071,838
;  EARLIER FILING DATE: 1998-05-01
;  NUMBER OF SEQ ID NOS: 324
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 143
;   LENGTH: 4
;   TYPE: PRT
;   ORGANISM: Arabidopsis sp.
US-09-177-249-143


  Query Match             42.2%;  Score 19;  DB 4;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 FYF 7
            |||
Db      2 FYF 4


RESULT   9
US-08-282-980B-2
; Sequence 2, Application US/08282980B
; Patent No. 5932189
;  GENERAL INFORMATION:
;    APPLICANT:  Dean, Richard T.
;    APPLICANT:  McBride, William
;    APPLICANT:  Lister-James, John
;    TITLE OF INVENTION:  Peptides
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McDonnell Boehnen Hulbert & Berghoff
;      STREET:  300 South Wacker Drive, Seventh Floor
;      CITY:  Chicago
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/282,980B
;      FILING DATE:  29-JUL-1994
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5932189nan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  92,385-I
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-913-0001
;      TELEFAX:  312-913-0002
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1..6
;      OTHER INFORMATION:  /label= Cyclic
;      OTHER INFORMATION:  /note= "The amino terminus and carboxyl terminus
;      OTHER INFORMATION:  are linked by a covalent bond; the amino terminus
;      OTHER INFORMATION:  is substituted with a methyl group."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1..6
;      OTHER INFORMATION:  /label= Variant residues
;      OTHER INFORMATION:  /note= "The Trp is in the D conformation; Xaa
;      OTHER INFORMATION:  is homocysteine."
US-08-282-980B-2


  Query Match             42.2%;  Score 19;  DB 2;  Length 6;
  Best Local Similarity   75.0%;  Pred. No. 1.7e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      5 FYFK 8
          ||:|
Db      1 FYWK 4


RESULT 10
US-08-282-980B-3
; Sequence 3, Application US/08282980B
; Patent No. 5932189
;  GENERAL INFORMATION:
;    APPLICANT:  Dean, Richard T.
;    APPLICANT:  McBride, William
;    APPLICANT:  Lister-James, John
;    TITLE OF INVENTION:  Peptides
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McDonnell Boehnen Hulbert & Berghoff
;      STREET:  300 South Wacker Drive, Seventh Floor
;      CITY:  Chicago
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/282,980B
;      FILING DATE:  29-JUL-1994
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5932189nan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  92,385-I
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-913-0001
;      TELEFAX:  312-913-0002
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1..6
;      OTHER INFORMATION:  /label= Cyclic
;      OTHER INFORMATION:  /note= "The amino terminus and carboxyl terminus
;      OTHER INFORMATION:  are linked by a covalent bond; the amino terminus
;      OTHER INFORMATION:  is substituted with a methyl group."
US-08-282-980B-3


  Query Match             42.2%;  Score 19;  DB 2;  Length 6;
  Best Local Similarity   75.0%;  Pred. No. 1.7e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 FYFK 8
          ||:|
Db      1 FYWK 4


RESULT 11
US-08-282-980B-4
; Sequence 4, Application US/08282980B
; Patent No. 5932189
;  GENERAL INFORMATION:
;    APPLICANT:  Dean, Richard T.
;    APPLICANT:  McBride, William
;    APPLICANT:  Lister-James, John
;    TITLE OF INVENTION:  Peptides
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McDonnell Boehnen Hulbert & Berghoff
;      STREET:  300 South Wacker Drive, Seventh Floor
;      CITY:  Chicago
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/282,980B
;      FILING DATE:  29-JUL-1994
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5932189nan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  92,385-I
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-913-0001
;      TELEFAX:  312-913-0002
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1..6
;      OTHER INFORMATION:  /label= Cyclic
;      OTHER INFORMATION:  /note= "The amino terminus and carboxyl terminus
;      OTHER INFORMATION:  are linked by a covalent bond; the amino terminus
;      OTHER INFORMATION:  is substituted with a methyl group."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1..6
;      OTHER INFORMATION:  /label= Variant residues
;      OTHER INFORMATION:  /note= "The Trp is in the D conformation; Xaa
;      OTHER INFORMATION:  is homocysteine."
US-08-282-980B-4


  Query Match             42.2%;  Score 19;  DB 2;  Length 6;
  Best Local Similarity   75.0%;  Pred. No. 1.7e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 FYFK 8
          ||:|
Db      1 FYWK 4


RESULT 12
US-08-482-228-170
; Sequence 170, Application US/08482228
; Patent No. 5968753
;  GENERAL INFORMATION:
;    APPLICANT:  Tseng-Law, Janet
;    APPLICANT:  Kobori, Joan A.
;    APPLICANT:  Al-Abdaly, Fahad A.
;    APPLICANT:  Guillermo, Roy
;    APPLICANT:  Helgerson, Sam L.
;    APPLICANT:  Deans, Robert J.
;    TITLE OF INVENTION:  POSITIVE AND POSITIVE/NEGATIVE CELL
;    TITLE OF INVENTION:  SELECTION MEDIATED BY PEPTIDE RELEASE
;    NUMBER OF SEQUENCES:  215
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Janice Guthrie, Ph.D.
```

```
;       STREET:  P.O. Box 15210
;       CITY:  Irvine
;       STATE:  California
;       COUNTRY:  USA
;       ZIP:  92713-5210
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/482,228
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Guthrie, Janice
;       REGISTRATION NUMBER:  35,170
;       REFERENCE/DOCKET NUMBER:  IT-4630CIP3
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (714) 440-5353
;       TELEFAX:  (714) 553-1952
;   INFORMATION FOR SEQ ID NO:  170:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  6 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-482-228-170


  Query Match             42.2%;  Score 19;  DB 2;  Length 6;
  Best Local Similarity   40.0%;  Pred. No. 1.7e+05;
  Matches    2;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 TFYFK 8
            :||::
Db       2 SFYYR 6


RESULT  13
US-08-931-095-2
; Sequence 2, Application US/08931095
; Patent No. 6017512
;  GENERAL INFORMATION:
;    APPLICANT:  Dean, Richard T.
;    APPLICANT:  McBride, William
;    APPLICANT:  Lister-James, John
;    TITLE OF INVENTION:  Radiolabeled Peptides
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McDonnell Boehnen Hulbert & Berghoff
;      STREET:  300 South Wacker Drive, 32nd Floor
;      CITY:  Chicago
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/931,095
;      FILING DATE:  15-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 6017512nan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  92,385-OO
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-913-0001
;      TELEFAX:  312-913-0002
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1..6
;      OTHER INFORMATION:  /label= Cyclic
;      OTHER INFORMATION:  /note= "The amino terminus and carboxyl terminus
;      OTHER INFORMATION:  are linked by a covalent bond; the amino terminus
;      OTHER INFORMATION:  is substituted with a methyl group."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1..3
;      OTHER INFORMATION:  /label= Variant residues
;      OTHER INFORMATION:  /note= "The Trp is in the D conformation; Xaa
;      OTHER INFORMATION:  is homohomocysteine.
US-08-931-095-2


  Query Match             42.2%;  Score 19;  DB 3;  Length 6;
  Best Local Similarity   75.0%;  Pred. No. 1.7e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 FYFK 8
            |||:|
Db       1 FYWK 4


RESULT  14
US-08-931-095-7
; Sequence 7, Application US/08931095
; Patent No. 6017512
;  GENERAL INFORMATION:
;    APPLICANT:  Dean, Richard T.
;    APPLICANT:  McBride, William
;    APPLICANT:  Lister-James, John
;    TITLE OF INVENTION:  Radiolabeled Peptides
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McDonnell Boehnen Hulbert & Berghoff
;      STREET:  300 South Wacker Drive, 32nd Floor
;      CITY:  Chicago
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/931,095
;      FILING DATE:  15-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 6017512nan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  92,385-OO
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-913-0001
;      TELEFAX:  312-913-0002
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Modified-site
```

```
;      LOCATION:  1..6
;      OTHER INFORMATION:  /label= Cyclic
;      OTHER INFORMATION:  /note= "The amino terminus and carboxyl terminus
;      OTHER INFORMATION:  are linked by a covalent bond; the amino terminus
;      OTHER INFORMATION:  is substituted with a methyl group."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1..3
;      OTHER INFORMATION:  /label= Variant residues
;      OTHER INFORMATION:  /note= "The Trp is in the D conformation."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  6
;      OTHER INFORMATION:  /label= Chelator
;      OTHER INFORMATION:  /note= "A BAT chelator is covalently linked
;      OTHER INFORMATION:  to the epsilon amino group of the Lys residue."
US-08-931-095-7


  Query Match             42.2%;  Score 19;  DB 3;  Length 6;
  Best Local Similarity   75.0%;  Pred. No. 1.7e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 FYFK 8
           ||:|
Db       1 FYWK 4


RESULT  15
US-08-931-095-9
; Sequence 9, Application US/08931095
; Patent No. 6017512
;  GENERAL INFORMATION:
;    APPLICANT:  Dean, Richard T.
;    APPLICANT:  McBride, William
;    APPLICANT:  Lister-James, John
;    TITLE OF INVENTION:  Radiolabeled Peptides
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McDonnell Boehnen Hulbert & Berghoff
;      STREET:  300 South Wacker Drive, 32nd Floor
;      CITY:  Chicago
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/931,095
;      FILING DATE:  15-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 6017512nan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  92,385-OO
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-913-0001
;      TELEFAX:  312-913-0002
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1..6
;      OTHER INFORMATION:  /label= Cyclic
;      OTHER INFORMATION:  /note= "The amino terminus and carboxyl terminus
;      OTHER INFORMATION:  are linked by a covalent bond; the amino terminus
;      OTHER INFORMATION:  is substituted with a methyl group."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1..3
;      OTHER INFORMATION:  /label= Variant residues
;      OTHER INFORMATION:  /note= "The Trp is in the D conformation."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  7
;      OTHER INFORMATION:  /label= Chelator
;      OTHER INFORMATION:  /note= "A BAM chelator is covalently linked
;      OTHER INFORMATION:  to the sidechain carboxyl group of Glu."
US-08-931-095-9


  Query Match             42.2%;  Score 19;  DB 3;  Length 6;
  Best Local Similarity   75.0%;  Pred. No. 1.7e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 FYFK 8
           ||:|
Db       1 FYWK 4


RESULT  16
US-08-482-528-170
; Sequence 170, Application US/08482528
; Patent No. 6017719
;  GENERAL INFORMATION:
;    APPLICANT:  Tseng-Law, Janet
;    APPLICANT:  Kobori, Joan A.
;    APPLICANT:  Al-Abdaly, Fahad A.
;    APPLICANT:  Guillermo, Roy
;    APPLICANT:  Helgerson, Sam L.
;    APPLICANT:  Deans, Robert J.
;    TITLE OF INVENTION:  POSITIVE AND POSITIVE/NEGATIVE CELL
;    TITLE OF INVENTION:  SELECTION MEDIATED BY PEPTIDE RELEASE
;    NUMBER OF SEQUENCES:  215
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Janice Guthrie, Ph.D.
;      STREET:  P.O. Box 15210
;      CITY:  Irvine
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92713-5210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/482,528
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Guthrie, Janice
;      REGISTRATION NUMBER:  35,170
;      REFERENCE/DOCKET NUMBER:  IT-4630CIP4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (714) 440-5353
;      TELEFAX:  (714) 553-1952
;  INFORMATION FOR SEQ ID NO:  170:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-482-528-170


  Query Match             42.2%;  Score 19;  DB 3;  Length 6;
  Best Local Similarity   40.0%;  Pred. No. 1.7e+05;
```

```
  Matches    2;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 TFYFK 8
          :||::
Db      2 SFYYR 6


RESULT  17
US-08-290-448A-82
; Sequence 82, Application US/08290448A
; Patent No. 5676954
;  GENERAL INFORMATION:
;    APPLICANT:  Rogers, Bruce
;    APPLICANT:  Klapper, David G.
;    APPLICANT:  Rafnar, Thorunn
;    APPLICANT:  Kuo, Mei-chang
;    TITLE OF INVENTION:  Allergenic Proteins From Ragweed and Uses
;    NUMBER OF SEQUENCES:  93
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street, suite 510
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109-1875
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/290,448A
;      FILING DATE:  August 15, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/529,951
;      FILING DATE:  May 29, 1990
;      APPLICATION NUMBER:  US 07/325,365
;      FILING DATE:  March 17, 1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Amy E. Mandragouras
;      REGISTRATION NUMBER:  36,207
;      REFERENCE/DOCKET NUMBER:  IMI-018CN
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)227-5941
;  INFORMATION FOR SEQ ID NO:  82:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-290-448A-82


  Query Match             42.2%;  Score 19;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 FYF 7
          |||
Db      5 FYF 7


RESULT  18
US-08-290-448A-82
; Sequence 82, Application US/08290448A
; Patent No. 5698204
;  GENERAL INFORMATION:
;    APPLICANT:  Rogers, Bruce
;    APPLICANT:  Klapper, David G.
;    APPLICANT:  Rafnar, Thorunn
;    APPLICANT:  Kuo, Mei-chang
;    TITLE OF INVENTION:  Allergenic Proteins From Ragweed and Uses
;    NUMBER OF SEQUENCES:  93
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street, suite 510
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109-1875
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/290,448A
;      FILING DATE:  August 15, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/529,951
;      FILING DATE:  May 29, 1990
;      APPLICATION NUMBER:  US 07/325,365
;      FILING DATE:  March 17, 1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Amy E. Mandragouras
;      REGISTRATION NUMBER:  36,207
;      REFERENCE/DOCKET NUMBER:  IMI-018CN
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)227-5941
;  INFORMATION FOR SEQ ID NO:  82:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-290-448A-82


  Query Match             42.2%;  Score 19;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 FYF 7
          |||
Db      5 FYF 7


RESULT  19
US-08-487-890A-81
; Sequence 81, Application US/08487890A
; Patent No. 5708149
;  GENERAL INFORMATION:
;    APPLICANT:  Loosmore, Sheena
;    APPLICANT:  Harkness, Robin
;    APPLICANT:  Schryvers, Anthony
;    APPLICANT:  Chong, Pele
;    APPLICANT:  Gray-Owen, Scott
;    APPLICANT:  Yang, Yan-Ping
;    APPLICANT:  Murdin, Andrew
;    APPLICANT:  Klein, Michel
;    TITLE OF INVENTION:  Transferrin Receptor Genes
;    NUMBER OF SEQUENCES:  147
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sim & McBurney
;      STREET:  6th Floor, 330 Unviersity Avenue
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5G 1R7
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,890A
;      FILING DATE:  07-JUN-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/175,116
;      FILING DATE:  29-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/148,968
;      FILING DATE:  08-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stewart, Michael I
;      REGISTRATION NUMBER:  24,973
;      REFERENCE/DOCKET NUMBER:  1038-466 MIS:jb
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 595-1155
;      TELEFAX:  (416) 595-1163
;  INFORMATION FOR SEQ ID NO:  81:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-487-890A-81


  Query Match             42.2%;  Score 19;  DB 1;  Length 7;
  Best Local Similarity   60.0%;  Pred. No. 1.7e+05;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 NATFY 6
           |  ||
Db       2 NGAFY 6


RESULT  20
US-08-175-069A-82
; Sequence 82, Application US/08175069A
; Patent No. 5776761
;  GENERAL INFORMATION:
;    APPLICANT:  Rogers, Bruce
;    APPLICANT:  Klapper, David G.
;    APPLICANT:  Rafnar, Thorunn
;    APPLICANT:  Kuo, Mei-chang
;    TITLE OF INVENTION:  Allergenic Proteins From Ragweed and Uses
;    NUMBER OF SEQUENCES:  93
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD, LLP
;      STREET:  60 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109-1875
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/175,069A
;      FILING DATE:  December 29, 1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/529,951
;      FILING DATE:  May 29, 1990
;      APPLICATION NUMBER:  US 07/325,365
;      FILING DATE:  March 17, 1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Amy E. Mandragouras
;      REGISTRATION NUMBER:  36,207
;      REFERENCE/DOCKET NUMBER:  IMI-018DV
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)227-5941
;  INFORMATION FOR SEQ ID NO:  82:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-175-069A-82


  Query Match             42.2%;  Score 19;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 FYF 7
           |||
Db       5 FYF 7


Search completed: July  8, 2002, 11:38:46
Job time: 53 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2002, 11:41:19 ; Search time 24.98 Seconds
(without alignments)
46.160 Million cell updates/sec

Title: US-09-461-061A-2
Perfect score: 60
Sequence: 1 TLTHTITKLNAE 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 1579

Minimum DB seq length: 0
Maximum DB seq length: 12

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	22	36.7	11	2	T17081	cytochrome-c oxida
2	21	35.0	8	2	T14906	hypothetical prote
3	21	35.0	10	2	D46285	formaldehyde dehyd
4	19	31.7	10	2	A39745	endo-glucosylceram
5	17	28.3	11	2	PQ0733	unidentified 6.0/1
6	17	28.3	12	2	PC4377	telomeric and tetr
7	17	28.3	12	2	F84132	hypothetical prote
8	17	28.3	12	4	S49073	frame shifted cyta
9	16	26.7	10	2	T13976	cytochrome-c oxida
10	16	26.7	10	2	T14215	cytochrome-c oxida
11	16	26.7	10	2	T14223	cytochrome-c oxida
12	16	26.7	10	2	A59173	nuclease Bh1 (EC 3
13	15	25.0	7	2	S29735	polyphosphate--glu
14	15	25.0	9	2	JQ1202	leader peptide - P
15	15	25.0	9	2	PH0917	T-cell receptor be
16	15	25.0	9	2	G85802	hypothetical prote
17	15	25.0	10	2	S63478	dihydrolipoamide d
18	15	25.0	11	2	S58244	pyrroloquinoline q
19	15	25.0	11	2	PC2330	cycloinulooligosac
20	15	25.0	12	2	B44818	extracellular lipa
21	15	25.0	12	2	PN0162	malate dehydrogena
22	15	25.0	12	2	I77529	estrogen receptor
23	15	25.0	12	2	PT0255	Ig heavy chain CRD
24	14	23.3	7	2	A34026	acetylcholinestera
25	14	23.3	7	2	PX0008	glucuronosyltransf
26	14	23.3	8	2	A47618	beta-galactosidase
27	14	23.3	9	1	AKLQIM	locustamyoinhibiti
28	14	23.3	9	2	S36898	ribosomal protein
29	14	23.3	9	2	PT0231	Ig heavy chain CDR
30	14	23.3	10	2	S23307	neurokinin A - rai
31	14	23.3	10	2	S23186	neurokinin A - Atl
32	14	23.3	10	2	B33143	pneumadin - human
33	14	23.3	10	2	A33143	pneumadin - rat
34	14	23.3	10	2	S59625	beta-galactosidase
35	14	23.3	10	2	D61440	polygalacturonase
36	14	23.3	10	2	A61007	hementin (EC 3.4.-
37	14	23.3	10	2	T14219	cytochrome-c oxida
38	14	23.3	11	2	PN0167	ribosomal protein
39	14	23.3	11	2	PU0034	dextransucrase (EC
40	14	23.3	12	2	S25485	transcription fact
41	14	23.3	12	2	PN0160	ribosomal protein
42	14	23.3	12	2	F61308	hemocyanin chain 5
43	14	23.3	12	2	C20907	Ig kappa-1 chain J
44	14	23.3	12	2	F20907	Ig kappa-2 chain J
45	14	23.3	12	2	I57678	gene rPLP-A protei

ALIGNMENTS

RESULT 1
T17081
cytochrome-c oxidase (EC 1.9.3.1) chain I - Phrynocephalus raddei mitochondrion (frag
C;Species: mitochondrion Phrynocephalus raddei
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 22-Oct-1999
C;Accession: T17081
R;Macey, J.R.; Larson, A.; Ananjeva, N.B.; Papenfuss, T.J.
J. Mol. Evol. 44, 660-674, 1997
A;Title: Evolutionary shifts in three major structural features of the mitochondrial
A;Reference number: Z18674; MUID:97315309
A;Accession: T17081
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-11 <MAC>
A;Cross-references: EMBL:U82691; NID:g3603148; PID:g3603151; PIDN:AAC62302.1
C;Genetics:
A;Genome: mitochondrion
A;Note: COI
C;Keywords: mitochondrion; oxidoreductase

Query Match 36.7%; Score 22; DB 2; Length 11;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 TLTH 4
          ||||
Db      3 TLTH 6
```

RESULT 2
T14906
hypothetical protein - parsley
C;Species: Petroselinum crispum (parsley)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T14906
R;Feldbrugge, M.; Sprenger, M.; Dinkelbach, M.; Yazaki, K.; Harter, K.; Weisshaar, B.
Plant Cell 6, 1607-1621, 1994
A;Title: Functional analysis of a light-responsive plant bZIP transcriptional regulat
A;Reference number: Z18259; MUID:95128172
A;Accession: T14906
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-8 <FEL>
A;Cross-references: EMBL:S75395; NID:g913201; PID:e194245

Query Match 35.0%; Score 21; DB 2; Length 8;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy      2 LTHTIT 7
```

```
          : ||:|
Db      1 MKHTLT 6


RESULT   3
D46285
formaldehyde dehydrogenase (glutathione) (EC 1.2.1.1) class III low activity form - Atla
C;Species: Gadus morhua (Atlantic cod)
C;Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 20-Mar-1996
C;Accession: D46285
R;Danielsson, O.; Jornvall, H.
Proc. Natl. Acad. Sci. U.S.A. 89, 9247-9251, 1992
A;Title: "Enzymogenesis": classical liver alcohol dehydrogenase origin from the glutathi
A;Reference number: A46285; MUID:93028441
A;Accession: D46285
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-10 <DAN>
A;Note: sequence extracted from NCBI backbone (NCBIP:116272)
C;Keywords: NAD; oxidoreductase


  Query Match             35.0%;  Score 21;  DB 2;  Length 10;
  Best Local Similarity   60.0%;  Pred. No. 7.5e+02;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 LTHTI 6
          :|||:
Db      4 VTHTL 8


RESULT   4
A39745
endo-glucosylceramidase activator II - Rhodococcus sp. (fragment)
C;Species: Rhodococcus sp.
C;Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 30-Sep-1993
C;Accession: A39745
R;Ito, M.; Ikegami, Y.; Yamagata, T.
J. Biol. Chem. 266, 7919-7926, 1991
A;Title: Activator proteins for glycosphingolipid hydrolysis by endoglycoceramidases. El
ble using these activator proteins.
A;Reference number: A39745; MUID:91210321
A;Accession: A39745
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-10 <ITO>


  Query Match             31.7%;  Score 19;  DB 2;  Length 10;
  Best Local Similarity   66.7%;  Pred. No. 1.8e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      5 TITKLN 10
          |||  |
Db      3 TITPFN 8


RESULT   5
PQ0733
unidentified 6.0/15K protein [imported] - rice (fragment)
C;Species: Oryza sativa (rice)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C;Accession: PQ0733
R;Komatsu, S.; Kajiwara, H.; Hirano, H.
Theor. Appl. Genet. 86, 935-942, 1993
A;Title: A rice protein library; a data-file of rice proteins separated by two-dimension
A;Reference number: PQ0696
A;Accession: PQ0733
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-11 <KOM>


  Query Match             28.3%;  Score 17;  DB 2;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 4.8e+03;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 HTI 6
          |||
Db      1 HTI 3


RESULT   6
PC4377
telomeric and tetraplex DNA binding protein qTBP42 VIII - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 28-Oct-1997 #sequence_revision 28-Oct-1997 #text_change 07-May-1999
C;Accession: PC4377
R;Sarig, G.; Weisman-Shomer, P.; Fry, M.
Biochem. Biophys. Res. Commun. 237, 617-623, 1997
A;Title: Telomeric and tetraplex DNA binding properties of qTBP42: A homologue of the
A;Reference number: PC4371; MUID:97445086
A;Accession: PC4377
A;Molecule type: protein
A;Residues: 1-12 <SAR>
C;Comment: This protein binds either strand of the telomeric DNA as well as unimolecu


  Query Match             28.3%;  Score 17;  DB 2;  Length 12;
  Best Local Similarity   50.0%;  Pred. No. 5.3e+03;
  Matches    2;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 HTIT 7
          ||::
Db      5 HTVS 8


RESULT   7
F84132
hypothetical protein BH3862 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: F84132
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: F84132
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-12 <STO>
A;Cross-references: GB:AP001520; GB:BA000004; NID:g10176401; PIDN:BAB07581.1; GSPDB:G
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH3862


  Query Match             28.3%;  Score 17;  DB 2;  Length 12;
  Best Local Similarity   25.0%;  Pred. No. 5.3e+03;
  Matches    2;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 LTHTITKL 9
          : : :||:
Db      1 MNYLLTKI 8


RESULT   8
S49073
frame shifted cytadherence accessory protein HMW3 - Mycoplasma pneumoniae (fragment)
N;Alternate names: H+-transporting ATP synthase (EC 3.6.1.34) alpha chain (misidentif
C;Species: Mycoplasma pneumoniae
C;Date: 28-Aug-1998 #sequence_revision 28-Aug-1998 #text_change 20-Apr-2000
C;Accession: S49073
R;Proft, T.; Herrmann, R.
Mol. Microbiol. 13, 337-348, 1994
A;Title: Identification and characterization of hitherto unknown Mycoplasma pneumonia
```

A;Reference number: S49059; MUID:95075318
A;Accession: S49073
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-12 <PRO>
A;Cross-references: EMBL:Z32665; NID:g474163; PIDN:CAA83583.1; PID:g581348
A;Experimental source: clone D2-16
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1994
A;Note: the GenBank entry MPDASA1, release 106.0, translates the ATC codon, 1 in this re
C;Comment: The nucleotide sequence translated in an alternative reading frame is identic
C;Genetics:
A;Genetic code: SGC3

Query Match 28.3%; Score 17; DB 4; Length 12;
Best Local Similarity 80.0%; Pred. No. 5.3e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TLTHT 5
||| |
Db 3 TLTMT 7

RESULT 9
T13976
cytochrome-c oxidase (EC 1.9.3.1) chain I - Cnemidophorus tigris mitochondrion (fragment
C;Species: mitochondrion Cnemidophorus tigris
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 11-May-2000
C;Accession: T13976
R;Macey, J.R.; Larson, A.; Ananjeva, N.B.; Fang, Z.; Papenfuss, T.J.
Mol. Biol. Evol. 14, 91-104, 1997
A;Title: Two novel gene orders and the role of light-strand replication in rearrangement
A;Reference number: Z17789; MUID:97153826
A;Accession: T13976
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-10 <MAC>
A;Cross-references: EMBL:U71332; NID:g1753236; PID:g1753239; PIDN:AAB48274.1
C;Genetics:
A;Genome: mitochondrion
A;Note: CO1
C;Keywords: mitochondrion; oxidoreductase

Query Match 26.7%; Score 16; DB 2; Length 10;
Best Local Similarity 75.0%; Pred. No. 6.7e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TITK 8
|||:
Db 2 TITR 5

RESULT 10
T14215
cytochrome-c oxidase (EC 1.9.3.1) chain I - Varanus griseu mitochondrion (fragment)
C;Species: mitochondrion Varanus griseu
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 11-May-2000
C;Accession: T14215
R;Macey, J.R.; Larson, A.; Ananjeva, N.B.; Fang, Z.; Papenfuss, T.J.
Mol. Biol. Evol. 14, 91-104, 1997
A;Title: Two novel gene orders and the role of light-strand replication in rearrangement
A;Reference number: Z17789; MUID:97153826
A;Accession: T14215
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-10 <MAC>
A;Cross-references: EMBL:U71334; NID:g3688056; PID:g1753271; PIDN:AAD12669.1
C;Genetics:
A;Genome: mitochondrion
A;Note: CO1
C;Keywords: mitochondrion; oxidoreductase

Query Match 26.7%; Score 16; DB 2; Length 10;
Best Local Similarity 75.0%; Pred. No. 6.7e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TITK 8
|||:
Db 2 TITR 5

RESULT 11
T14223
cytochrome-c oxidase (EC 1.9.3.1) chain I - Xantusia vigilis mitochondrion (fragment)
C;Species: mitochondrion Xantusia vigilis
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 11-May-2000
C;Accession: T14223
R;Macey, J.R.; Larson, A.; Ananjeva, N.B.; Fang, Z.; Papenfuss, T.J.
Mol. Biol. Evol. 14, 91-104, 1997
A;Title: Two novel gene orders and the role of light-strand replication in rearrangem
A;Reference number: Z17789; MUID:97153826
A;Accession: T14223
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-10 <MAC>
A;Cross-references: EMBL:U71328; NID:g1753276; PID:g1753279; PIDN:AAB48291.1
C;Genetics:
A;Genome: mitochondrion
A;Note: CO1
C;Keywords: mitochondrion; oxidoreductase

Query Match 26.7%; Score 16; DB 2; Length 10;
Best Local Similarity 75.0%; Pred. No. 6.7e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TITK 8
|||:
Db 2 TITR 5

RESULT 12
A59173
nuclease Bh1 (EC 3.1.-.-) - Basidiobolus haptosporus (fragment)
C;Species: Basidiobolus haptosporus
C;Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 18-Feb-2000
C;Accession: A59173
R;Desai, N.; Shankar, V.
submitted to the Protein Sequence Database, February 2000
A;Description: Single-strand-specific, guanylic acid preferential nuclease from Basid
A;Reference number: A59173
A;Accession: A59173
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-10 <DES>
A;Note: extracellular, single-strand-specific nuclease
C;Keywords: hydrolase

Query Match 26.7%; Score 16; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 6.7e+03;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LTHTIT 7
| | :|
Db 4 LGHLLT 9

RESULT 13
S29735
polyphosphate--glucose phosphotransferase (EC 2.7.1.63) - Propionibacterium freudenre
C;Species: Propionibacterium freudenreichii subsp. shermanii
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 07-May-1999
C;Accession: S29735

R;Phillips, N.F.B.; Horn, P.J.; Wood, H.G.
Arch. Biochem. Biophys. 300, 309-319, 1993
A;Title: The polyphosphate- and ATP-dependent glucokinase from Propionibacterium sherman
A;Reference number: S29735; MUID:93143332
A;Accession: S29735
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <PHI>
C;Keywords: phosphotransferase

Query Match 25.0%; Score 15; DB 2; Length 7;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy      3 THTI 6
          || :
Db      1 THVL 4
```

RESULT 14
JQ1202
leader peptide - Pseudomonas sp. plasmid R1033 transposon Tn1696
C;Species: Pseudomonas sp.
C;Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 12-Dec-1997
C;Accession: JQ1202
R;Stokes, H.W.; Hall, R.M.
Plasmid 26, 10-19, 1991
A;Title: Sequence analysis of the inducible chloramphenicol resistance determinant in th
A;Reference number: JQ1201; MUID:92052679
A;Accession: JQ1202
A;Molecule type: DNA
A;Residues: 1-9 <STO>
A;Cross-references: GB:M60454
C;Comment: This peptide is a potential translational attenuation signal for cmlA gene.
C;Genetics:
A;Genome: plasmid
C;Superfamily: unassigned leader peptides

Query Match 25.0%; Score 15; DB 2; Length 9;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy      7 TKLNAE 12
          |  ||:
Db      3 TSKNAD 8
```

RESULT 15
PH0917
T-cell receptor beta chain V-D-J region (isolate 3) - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 30-May-1997
C;Accession: PH0917
R;Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.
J. Exp. Med. 174, 1467-1476, 1991
A;Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergi
A;Reference number: PH0891; MUID:92078857
A;Accession: PH0917
A;Molecule type: mRNA
A;Residues: 1-9 <GOL>
A;Experimental source: concanavalin A-activated lymphoblast
C;Keywords: T-cell receptor

Query Match 25.0%; Score 15; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     10 NAE 12
          |||
Db      7 NAE 9
```

RESULT 16
G85802
hypothetical protein Z2947 [imported] - Escherichia coli (strain O157:H7, substrain E
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: G85802
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: G85802
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-9 <STO>
A;Cross-references: GB:AE005174; NID:g12515957; PIDN:AAG56883.1; GSPDB:GN00145; UWGP:
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z2947

Query Match 25.0%; Score 15; DB 2; Length 9;
Best Local Similarity 33.3%; Pred. No. 2.8e+05;
Matches 3; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```
Qy      2 LTHTITKLN 10
          :|:|    |
Db      1 MTYTFMLSN 9
```

RESULT 17
S63478
dihydrolipoamide dehydrogenase (EC 1.8.1.4) beta chain E1 - Pseudomonas aeruginosa (f
N;Alternate names: branched-chain-oxoacid dehydrogenase chain E1
C;Species: Pseudomonas aeruginosa
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
C;Accession: S63478
R;Hester, K.; Luo, J.; Burns, G.; Braswell, E.H.; Sokatch, J.R.
Eur. J. Biochem. 233, 828-836, 1995
A;Title: Purification of active E1-alpha(2)-beta(2) of Pseudomonas putida branched-ch
A;Reference number: S63475; MUID:96085147
A;Accession: S63478
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-10 <HES>
C;Keywords: FAD; flavoprotein; lipoamide; NAD; oxidoreductase

Query Match 25.0%; Score 15; DB 2; Length 10;
Best Local Similarity 40.0%; Pred. No. 1e+04;
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy      5 TITKL 9
          |:| :
Db      6 TVTSM 10
```

RESULT 18
S58244
pyrroloquinoline quinone synthesis C - Pseudomonas fluorescens (fragment)
C;Species: Pseudomonas fluorescens
C;Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 08-Oct-1999
C;Accession: S58244
R;Schnider, U.; Keel, C.; Defago, G.; Haas, D.
submitted to the EMBL Data Library, May 1995
A;Description: Tn5-directed cloning of pqq genes from Pseudomonas fluorescens CHA0: t
A;Reference number: S58239
A;Accession: S58244
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-11 <SCH>

A;Cross-references: EMBL:X87299; NID:g929799; PIDN:CAA60734.1; PID:g929806

Query Match 25.0%; Score 15; DB 2; Length 11;
Best Local Similarity 57.1%; Pred. No. 1.1e+04;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy      5 TITKLNA 11
          | | |:|
Db      2 TDTPLSA 8
```

RESULT 19
PC2330
cycloinulooligosaccharide fructanotransferase (EC 2.4.-.-) - Bacillus circulans (MCI-255
C;Species: Bacillus circulans
C;Date: 21-Mar-1995 #sequence_revision 26-May-1995 #text_change 07-May-1999
C;Accession: PC2330
R;Kushibe, S.; Mitsui, K.; Yamagishi, M.; Yamada, K.; Morimoto, Y.
Biosci. Biotechnol. Biochem. 59, 31-34, 1995
A;Title: Purification and characterization of cycloinulooligosaccharide fructanotransfer
A;Reference number: PC2330; MUID:95201377
A;Accession: PC2330
A;Molecule type: protein
A;Residues: 1-11 <KUS>
C;Comment: This enzyme hydrolyzes beta-(2-1) glycosidic linkages and acts in intermolecu
C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 25.0%; Score 15; DB 2; Length 11;
Best Local Similarity 22.2%; Pred. No. 1.1e+04;
Matches 2; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```
Qy      4 HTITKLNAE 12
          |   ::| :
Db      2 HLFYQMNPQ 10
```

RESULT 20
B44818
extracellular lipase - Pseudomonas aeruginosa (fragment)
C;Species: Pseudomonas aeruginosa
C;Date: 31-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 18-Nov-1994
C;Accession: B44818
R;Gilbert, E.J.; Cornish, A.; Jones, C.W.
J. Gen. Microbiol. 137, 2223-2229, 1991
A;Title: Purification and properties of extracellular lipase from Pseudomonas aeruginosa
A;Reference number: A44818; MUID:92085040
A;Accession: B44818
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-12 <GIL>
A;Experimental source: strain EF228
A;Note: sequence extracted from NCBI backbone (NCBIP:70393)

Query Match 25.0%; Score 15; DB 2; Length 12;
Best Local Similarity 50.0%; Pred. No. 1.3e+04;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy      3 THTITK 8
          |:| |:
Db      1 TYTQTQ 6
```

Search completed: July 8, 2002, 11:43:57
Job time: 158 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 8, 2002, 11:43:29 ; Search time 13.1 Seconds
(without alignments)
35.468 Million cell updates/sec

Title: US-09-461-061A-2
Perfect score: 60
Sequence: 1 TLTHTITKLNAE 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 463

Minimum DB seq length: 0
Maximum DB seq length: 12

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	16	26.7	12	1	HCYB_MEGCR	Q10584 megathura c
2	15	25.0	11	1	PQQC_PSEFL	P55173 pseudomonas
3	15	25.0	12	1	PPK4_PERFU	P82690 periplaneta
4	15	25.0	12	1	UH03_RAT	P56572 rattus norv
5	15	25.0	12	1	V25K_WSSV	P82004 white spot
6	14	23.3	8	1	FUSS_FUSSO	P81010 fusarium so
7	14	23.3	9	1	LMIP_LOCMI	P31799 locusta mig
8	14	23.3	10	1	MALE_KLEPN	Q05564 klebsiella
9	14	23.3	10	1	PNEU_HUMAN	P22103 homo sapien
10	14	23.3	10	1	PNEU_RAT	P21996 rattus norv
11	14	23.3	10	1	SPI_HALRO	Q10997 halocynthia
12	14	23.3	10	1	TKNB_ONCMY	P28500 oncorhynchu
13	14	23.3	10	1	UHA3_HUMAN	P40930 homo sapien
14	13	21.7	8	1	ACI_THUAL	P18691 thunnus alb
15	13	21.7	8	1	AKH_TABAT	P14595 tabanus atr
16	13	21.7	8	1	HTF2_PERAM	P04549 periplaneta
17	13	21.7	9	1	PPK1_PERAM	P82691 periplaneta
18	13	21.7	9	1	XYLA_STRSQ	P19149 streptomyce
19	13	21.7	10	1	AMPN_HELAM	P81731 helicoverpa
20	13	21.7	10	1	HTF2_CARMO	P11385 carausius m
21	13	21.7	10	1	HTF_TABAT	P14596 tabanus atr
22	13	21.7	12	1	PPK4_PERAM	P82619 periplaneta
23	12	20.0	9	1	DCML_PSECF	P19913 pseudomonas
24	12	20.0	10	1	GON2_CHEPR	P80678 chelyosoma
25	12	20.0	11	1	TKNA_RANRI	P29207 rana ridibu
26	12	20.0	11	1	TKND_RANCA	P22691 rana catesb
27	12	20.0	12	1	V23K_WSSV	P82005 white spot
28	11	18.3	5	1	BIOB_CITFR	P12997 citrobacter
29	11	18.3	6	1	TRPI_PSEPU	P36414 pseudomonas
30	11	18.3	7	1	IGAO_DACDE	P06294 dactylium d
31	11	18.3	8	1	AKH_MELML	P25423 melolontha
32	11	18.3	8	1	HTF_TENMO	P25419 tenebrio mo
33	11	18.3	8	1	PLP_BRANA	P81707 brassica na
34	11	18.3	8	1	RPCH_PANBO	P08939 pandalus bo
35	11	18.3	9	1	MOSH_CLYJA	P19852 clypeaster
36	11	18.3	9	1	PGLR_DIAAB	P81179 diaprepes a
37	11	18.3	9	1	PTSP_BOMMO	P82003 bombyx mori
38	11	18.3	9	1	TAL1_PICJA	P17440 pichia jadi
39	11	18.3	9	1	TAL3_PICJA	P17441 pichia jadi
40	11	18.3	10	1	COX4_THUOB	P80971 thunnus obe
41	11	18.3	10	1	COXA_ONCMY	P80328 oncorhynchu
42	11	18.3	10	1	FARP_MANSE	P18523 manduca sex
43	11	18.3	10	1	LABA_JATMU	P13270 jatropha mu
44	11	18.3	10	1	Q2OG_COMTE	P80466 comamonas t
45	11	18.3	10	1	TKNB_RANRI	P29135 rana ridibu

ALIGNMENTS

RESULT 1
HCYB_MEGCR
ID HCYB_MEGCR STANDARD; PRT; 12 AA.
AC Q10584;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hemocyanin B chain (KLH-B) (Fragment).
OS Megathura crenulata (Giant keyhole limpet).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Archaeogastropoda;
OC Fissurellidae; Megathura.
OX NCBI_TaxID=55429;
RN [1]
RP SEQUENCE.
RX MEDLINE=96208935; PubMed=8829804;
RA Swerdlow R.D., Ebert R.F., Lee P., Bonaventura C., Miller K.I.;
RT "Keyhole limpet hemocyanin: structural and functional
RT characterization of two different subunits and multimers.";
RL Comp. Biochem. Physiol. 113B:537-548(1996).
CC -!- FUNCTION: HEMOCYANINS ARE COPPER-CONTAINING OXYGEN CARRIERS
CC OCCURRING FREELY DISSOLVED IN THE HEMOLYMPH OF MANY MOLLUSKS AND
CC ARTHROPODS.
CC -!- SUBUNIT: DIDECAMERS AND EXTENDED MULTIMERS.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: HEMOLYMPH.
CC -!- BIOTECHNOLOGY: Potent immunogen used classically as a carrier
CC protein for haptens and more recently in human vaccines and for
CC immunotherapy of bladder cancer.
CC -!- SIMILARITY: BELONGS TO THE TYROSINASE FAMILY. HEMOCYANIN
CC SUBFAMILY.
DR InterPro; IPR000896; Hemocyanin.
DR PROSITE; PS00209; HEMOCYANIN_1; PARTIAL.
DR PROSITE; PS00210; HEMOCYANIN_2; PARTIAL.
KW Oxygen transport; Transport; Copper; Glycoprotein;
KW Hemolymph.
FT NON_TER 12 12
SQ SEQUENCE 12 AA; 1345 MW; CBFBEAA44A432412 CRC64;

Query Match 26.7%; Score 16; DB 1; Length 12;
Best Local Similarity 25.0%; Pred. No. 2.6e+03;
Matches 2; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```
Qy      5 TITKLNAE 12
          |: : | :
Db      3 TVVRKNVD 10
```

RESULT 2
PQQC_PSEFL
ID PQQC_PSEFL STANDARD; PRT; 11 AA.
AC P55173;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Coenzyme PQQ synthesis protein C (Fragment).

```
GN   PQQC.
OS   Pseudomonas fluorescens.
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Pseudomonas.
OX   NCBI_TaxID=294;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CHA0;
RX   MEDLINE=96064397; PubMed=8526497;
RA   Schnider U., Keel C., Defago G., Haas D.;
RT   "Tn5-directed cloning of pqq genes from Pseudomonas fluorescens CHA0:
RT   mutational inactivation of the genes results in overproduction of the
RT   antibiotic pyoluteorin.";
RL   Appl. Environ. Microbiol. 61:3856-3864(1995).
CC   -!- FUNCTION: REQUIRED FOR COENZYME PYRROLO-QUINOLINE-QUINONE (PQQ)
CC       BIOSYNTHESIS.
CC   -!- SIMILARITY: TO OTHER BACTERIAL PQQC.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X87299; CAA60734.1; -.
KW   PQQ.
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  1182 MW;  89DF46E4C5B73771 CRC64;


  Query Match             25.0%;  Score 15;  DB 1;  Length 11;
  Best Local Similarity   57.1%;  Pred. No. 3.7e+03;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 TITKLNA 11
            | | |:|
Db       2 TDTPLSA 8


RESULT   3
PPK4_PERFU
ID   PPK4_PERFU     STANDARD;      PRT;    12 AA.
AC   P82690;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Pyrokinin-4 (PEF-PK-4) (YXPRL-amide).
OS   Periplaneta fuliginosa (Smokybrown cockroach).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC   Blattoidea; Blattidae; Periplaneta.
OX   NCBI_TaxID=36977;
RN   [1]
RP   SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC   TISSUE=Corpora cardiaca;
RX   MEDLINE=20189894; PubMed=10723010;
RA   Predel R., Eckert M.;
RT   "Tagma-specific distribution of FXPRLamides in the nervous system of
RT   the American cockroach.";
RL   J. Comp. Neurol. 419:352-363(2000).
CC   -!- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
CC       (MYOTROPIC ACTIVITY).
CC   -!- MASS SPECTROMETRY: MW=1437.9; METHOD=MALDI.
CC   -!- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.
DR   InterPro; IPR001484; Pyrokinin.
DR   PROSITE; PS00539; PYROKININ; FALSE_NEG.
KW   Neuropeptide; Amidation; Pyrokinin.
FT   MOD_RES      12     12       AMIDATION.
SQ   SEQUENCE   12 AA;  1439 MW;  F9B53049FF42CAA1 CRC64;


  Query Match             25.0%;  Score 15;  DB 1;  Length 12;
  Best Local Similarity   40.0%;  Pred. No. 4e+03;
  Matches    2;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 LTHTI 6
            |:| :
Db       3 LSHDV 7


RESULT   4
UH03_RAT
ID   UH03_RAT       STANDARD;      PRT;    12 AA.
AC   P56572;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Unknown protein from 2D-page of heart tissue (Spot P3) (Fragment).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=WISTAR; TISSUE=Heart;
RA   Li X.-P., Pleissner K.-P., Scheler C., Regitz-Zagrosek V., Salikov J.,
RA   Jungblut P.R.;
RL   Submitted (SEP-1998) to the SWISS-PROT data bank.
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 8.3, ITS MW IS: 28 kDa.
FT   UNSURE       2      2
FT   UNSURE       9      9
FT   NON_TER     12     12
SQ   SEQUENCE   12 AA;  1255 MW;  46F58D101DC33053 CRC64;


  Query Match             25.0%;  Score 15;  DB 1;  Length 12;
  Best Local Similarity   50.0%;  Pred. No. 4e+03;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 HTITKL 9
            ||  |:
Db       3 HTKIKV 8


RESULT   5
V25K_WSSV
ID   V25K_WSSV      STANDARD;      PRT;    12 AA.
AC   P82004;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   25 kDa structural polyprotein (Fragment).
OS   White spot syndrome virus (WSSV).
OC   Viruses; Unassigned viruses.
OX   NCBI_TaxID=92652;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=South Carolina;
RX   MEDLINE=20214217; PubMed=10752552;
RA   Wang Q., Poulos B.T., Lightner D.V.;
RT   "Protein analysis of geographic isolates of shrimp white spot syndrome
RT   virus.";
RL   Arch. Virol. 145:263-274(2000).
CC   -!- FUNCTION: STRUCTURAL COMPONENT OF THE VIRION.
FT   NON_TER     12     12
SQ   SEQUENCE   12 AA;  1283 MW;  C5409AD9ECB731A9 CRC64;


  Query Match             25.0%;  Score 15;  DB 1;  Length 12;
  Best Local Similarity   30.0%;  Pred. No. 4e+03;
  Matches    3;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 LTHTITKLNA 11
```

```
          |: |:: : |
Db      3 LSFTLSVVTA 12


RESULT   6
FUSS_FUSSO
ID   FUSS_FUSSO     STANDARD;      PRT;     8 AA.
AC   P81010;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Allergen Fus s I3596* (Fragment).
OS   Fusarium solani (subsp. pisi) (Nectria haematococca).
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Hypocreales; Nectriaceae; Haematonectria.
OX   NCBI_TaxID=109625;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=IARI 3596; TISSUE=Mycelium;
RA   Verma J., Gangal S.V.;
RL   Submitted (JUL-1997) to the SWISS-PROT data bank.
KW   Allergen.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  898 MW;  C372C441F5B69041 CRC64;


  Query Match             23.3%;  Score 14;  DB 1;  Length 8;
  Best Local Similarity   20.0%;  Pred. No. 1e+05;
  Matches    1;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 LTHTI 6
          ::| :
Db      3 MSHNV 7


RESULT   7
LMIP_LOCMI
ID   LMIP_LOCMI     STANDARD;      PRT;     9 AA.
AC   P31799;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   01-OCT-1993 (Rel. 27, Last annotation update)
DE   Locustamyoinhibiting peptide (LOM-MIP).
OS   Locusta migratoria (Migratory locust).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC   Acridomorpha; Acridoidea; Acrididae; Locusta.
OX   NCBI_TaxID=7004;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92179466; PubMed=1796179;
RA   Schoofs L., Holman G.M., Hayes T.K., Nachman R.J., de Loof A.;
RT   "Isolation, identification and synthesis of locustamyoinhibiting
RT   peptide (LOM-MIP), a novel biologically active neuropeptide from
RT   Locusta migratoria.";
RL   Regul. Pept. 36:111-119(1991).
CC   -!- FUNCTION: SUPPRESSES SPONTANEOUS CONTRACTIONS OF THE HINDGUT AND
CC       OVIDUCT.
CC   -!- TISSUE SPECIFICITY: NEURONS LOCATED IN TWO VENTRAL CELL CLUSTERS
CC       IN THE SUBOESOPHAGEAL GANGLION.
DR   PIR; A60065; AKLQIM.
KW   Amidation; Neuropeptide.
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  1060 MW;  387D7DD4472AB6C3 CRC64;


  Query Match             23.3%;  Score 14;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      9 LNA 11
          |||
Db      5 LNA 7


RESULT   8
MALE_KLEPN
ID   MALE_KLEPN     STANDARD;      PRT;    10 AA.
AC   Q05564;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Maltose-binding periplasmic protein (Maltodextrin-binding protein)
DE   (MMBP) (Fragment).
GN   MALE.
OS   Klebsiella pneumoniae.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Klebsiella.
OX   NCBI_TaxID=573;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1033-5P14 / KAY2026;
RX   MEDLINE=93211295; PubMed=8459773;
RA   Bachellier S., Perrin D., Hofnung M., Gilson E.;
RT   "Bacterial interspersed mosaic elements (BIMEs) are present in the
RT   genome of Klebsiella.";
RL   Mol. Microbiol. 7:537-544(1993).
CC   -!- FUNCTION: MALE IS INVOLVED IN THE HIGH-AFFINITY MALTOSE MEMBRANE
CC       TRANSPORT SYSTEM. INITIAL RECEPTOR FOR THE ACTIVE TRANSPORT OF AND
CC       CHEMOTAXIS TOWARD MALTOOLIGOSACCHARIDES.
CC   -!- SUBCELLULAR LOCATION: Periplasmic.
CC   -!- SIMILARITY: BELONGS TO THE BACTERIAL EXTRACELLULAR SOLUTE-BINDING
CC       PROTEIN FAMILY 1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X68329; CAA48406.1; -.
DR   InterPro; IPR000567; SBP_bac_1.
DR   PROSITE; PS01037; SBP_BACTERIAL_1; PARTIAL.
KW   Transport; Sugar transport; Periplasmic.
FT   NON_TER       1      1
SQ   SEQUENCE   10 AA;  1159 MW;  8FD8DC4415A6DDDA CRC64;


  Query Match             23.3%;  Score 14;  DB 1;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 5.1e+03;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      6 ITK 8
          |||
Db      8 ITK 10


RESULT   9
PNEU_HUMAN
ID   PNEU_HUMAN     STANDARD;      PRT;    10 AA.
AC   P22103;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Pneumadin (PNM).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Lung;
RX   MEDLINE=91110910; PubMed=2274681;
RA   Batra V.K., Mathur M., Mir S.A., Kapoor R., Kumar M.A.;
```

```
RT   "Pneumadin: a new lung peptide which triggers antidiuresis.";
RL   Regul. Pept. 30:77-87(1990).
CC   -!- FUNCTION: ANTIDIURETIC PEPTIDE THAT TRIGGERS THE RELEASE OF ADH.
DR   PIR; B33143; B33143.
KW   Amidation.
FT   MOD_RES      10     10       AMIDATION.
SQ   SEQUENCE   10 AA;  956 MW;  640378DAA723276B CRC64;


  Query Match             23.3%;  Score 14;  DB 1;  Length 10;
  Best Local Similarity   75.0%;  Pred. No. 5.1e+03;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      8 KLNA 11
          ||:|
Db      5 KLDA 8


RESULT  10
PNEU_RAT
ID   PNEU_RAT       STANDARD;      PRT;    10 AA.
AC   P21996;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Pneumadin (PNM).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE, AND SYNTHESIS.
RC   TISSUE=Lung;
RX   MEDLINE=91110910; PubMed=2274681;
RA   Batra V.K., Mathur M., Mir S.A., Kapoor R., Kumar M.A.;
RT   "Pneumadin: a new lung peptide which triggers antidiuresis.";
RL   Regul. Pept. 30:77-87(1990).
CC   -!- FUNCTION: THIS ANTIDIURETIC PEPTIDE TRIGGERS THE RELEASE OF ADH.
DR   PIR; A33143; A33143.
KW   Amidation.
FT   MOD_RES      10     10       AMIDATION.
SQ   SEQUENCE   10 AA;  1048 MW;  641D00DAA723276B CRC64;


  Query Match             23.3%;  Score 14;  DB 1;  Length 10;
  Best Local Similarity   75.0%;  Pred. No. 5.1e+03;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      8 KLNA 11
          ||:|
Db      5 KLDA 8


RESULT  11
SPI_HALRO
ID   SPI_HALRO      STANDARD;      PRT;    10 AA.
AC   Q10997;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Serine proteinase inhibitor (Fragment).
OS   Halocynthia roretzi (Sea squirt).
OC   Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC   Stolidobranchia; Pyuridae; Halocynthia.
OX   NCBI_TaxID=7729;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Hemolymph;
RX   MEDLINE=96321313; PubMed=8759295;
RA   Shishikura F., Abe T., Ohtake S.-I., Tanaka K.;
RT   "Purification and characterization of a 58,000-Da proteinase
RT   inhibitor from the hemolymph of a solitary ascidian, Halocynthia
RT   roretzi.";
RL   Comp. Biochem. Physiol. 114B:1-9(1996).
CC   -!- FUNCTION: STRONGLY INHIBITS TRYPSIN AND PLASMA ENZYME(S) ACTIVITY.
CC   -!- SUBUNIT: MONOMER.
CC   -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR   InterPro; IPR000215; Serpin.
DR   PROSITE; PS00284; SERPIN; PARTIAL.
KW   Serpin; Serine protease inhibitor; Glycoprotein; Plasma.
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1104 MW;  4225C73B1B187AA3 CRC64;


  Query Match             23.3%;  Score 14;  DB 1;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 5.1e+03;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      7 TKLNAE 12
          || : |
Db      1 TKKDGE 6


RESULT  12
TKNB_ONCMY
ID   TKNB_ONCMY     STANDARD;      PRT;    10 AA.
AC   P28500;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Neurokinin A (Substance K) (Neuromedin L).
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri), and
OS   Gadus morhua (Atlantic cod).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=8022, 8049;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Brain;
RX   MEDLINE=92298992; PubMed=1376687;
RA   Jensen J., Conlon J.M.;
RT   "Substance-P-related and neurokinin-A-related peptides from the brain
RT   of the cod and trout.";
RL   Eur. J. Biochem. 206:659-664(1992).
CC   -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC       EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC       SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC       MUSCLES.
CC   -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
DR   PIR; S23186; S23186.
DR   InterPro; IPR002040; Tachykinin.
DR   PROSITE; PS00267; TACHYKININ; 1.
KW   Tachykinin; Neuropeptide; Amidation.
FT   MOD_RES      10     10       AMIDATION (BY SIMILARITY).
SQ   SEQUENCE   10 AA;  1145 MW;  136B4062C9D5B440 CRC64;


  Query Match             23.3%;  Score 14;  DB 1;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 5.1e+03;
  Matches    2;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      8 KLNA 11
          |:|:
Db      2 KINS 5


RESULT  13
UHA3_HUMAN
ID   UHA3_HUMAN     STANDARD;      PRT;    10 AA.
AC   P40930;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Unknown protein from 2D-page of heart (Spot 7513) (Fragment).
OS   Homo sapiens (Human).
```

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Heart;
RX   MEDLINE=95203287; PubMed=7895732;
RA   Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;
RT   "The human myocardial two-dimensional gel protein database: update
RT   1994.";
RL   Electrophoresis 15:1459-1465(1994).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 6.8, ITS MW IS: 47.3 kDa.
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1049 MW;  6BBCDE41A041B76B CRC64;


  Query Match             23.3%;  Score 14;  DB 1;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 5.1e+03;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 TIT 7
          |||
Db      5 TIT 7


RESULT  14
ACI_THUAL
ID   ACI_THUAL      STANDARD;      PRT;     8 AA.
AC   P18691;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-NOV-1990 (Rel. 16, Last annotation update)
DE   Angiotensin-converting enzyme inhibitor.
OS   Thunnus albacares (Yellowfin tuna) (Neothunnus macropterus).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
OC   Scombridae; Thunnus.
OX   NCBI_TaxID=8236;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Muscle;
RX   MEDLINE=88326322; PubMed=3415688;
RA   Kohama Y., Matsumoto S., Oka H., Teramoto T., Okabe M., Mimura T.;
RT   "Isolation of angiotensin-converting enzyme inhibitor from tuna
RT   muscle.";
RL   Biochem. Biophys. Res. Commun. 155:332-337(1988).
DR   PIR; A31570; A31570.
SQ   SEQUENCE   8 AA;  953 MW;  6AA863733051F1B7 CRC64;


  Query Match             21.7%;  Score 13;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 TH 4
          ||
Db      2 TH 3


RESULT  15
AKH_TABAT
ID   AKH_TABAT      STANDARD;      PRT;     8 AA.
AC   P14595;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   Adipokinetic hormone (AKH) (Dipteran corpora cardiaca factor I)
DE   (DCC I).
OS   Tabanus atratus (Horse fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Tabanomorpha;
OC   Tabanidae; Tabanus.
OX   NCBI_TaxID=7207;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Corpora cardiaca;
RX   MEDLINE=90046758; PubMed=2813385;
RA   Jaffe H., Raina A.K., Riley C.T., Fraser B.A., Nachman R.J.,
RA   Vogel V.W., Zhang Y.-S., Hayes D.K.;
RT   "Primary structure of two neuropeptide hormones with adipokinetic and
RT   hypotrehalosemic activity isolated from the corpora cardiaca of horse
RT   flies (Diptera).";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:8161-8164(1989).
CC   -!- FUNCTION: THIS HORMONE, RELEASED FROM CELLS IN THE CORPORA
CC       CARDIACA AFTER THE BEGINNING OF FLIGHT, CAUSES RELEASE OF
CC       DIGLYCERIDES FROM THE FAT BODY AND THEN STIMULATES THE FLIGHT
CC       MUSCLES TO USE THESE DIGLYCERIDES AS AN ENERGY SOURCE.
CC   -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR   PIR; A33995; A33995.
DR   InterPro; IPR002047; AKH.
DR   PROSITE; PS00256; AKH; 1.
KW   Neuropeptide; Amidation; Flight.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  949 MW;  86786771A9D1A736 CRC64;


  Query Match             21.7%;  Score 13;  DB 1;  Length 8;
  Best Local Similarity   75.0%;  Pred. No. 1e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 LTHT 5
          || |
Db      2 LTFT 5


RESULT  16
HTF2_PERAM
ID   HTF2_PERAM     STANDARD;      PRT;     8 AA.
AC   P04549;
DT   13-AUG-1987 (Rel. 05, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   Hypertrehalosaemic factor II (Neuropeptide M-II) (Periplanetin CC-2)
DE   (PeA-CAH-II) (LeD-CC-II) (Hypertrehalosaemic neuropeptide II).
OS   Periplaneta americana (American cockroach),
OS   Leptinotarsa decemlineata (Colorado potato beetle), and
OS   Blatta orientalis (Oriental cockroach).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC   Blattoidea; Blattidae; Periplaneta.
OX   NCBI_TaxID=6978, 7539, 6976;
RN   [1]
RP   SEQUENCE.
RC   SPECIES=P.americana;
RX   MEDLINE=85046530; PubMed=6548628;
RA   Witten J.L., Schaffer M.H., O'Shea M., Cook J.C., Hemling M.E.,
RA   Rinehart K.L. Jr.;
RT   "Structures of two cockroach neuropeptides assigned by fast atom
RT   bombardment mass spectrometry.";
RL   Biochem. Biophys. Res. Commun. 124:350-358(1984).
RN   [2]
RP   SEQUENCE.
RC   SPECIES=P.americana;
RX   MEDLINE=84298179; PubMed=6591205;
RA   Scarborough R.M., Jamieson G.C., Kalish F., Kramer S.J., McEnroe G.A.,
RA   Miller C.A., Schooley D.A.;
RT   "Isolation and primary structure of two peptides with
RT   cardioacceleratory and hyperglycemic activity from the corpora
RT   cardiaca of Periplaneta americana.";
RL   Proc. Natl. Acad. Sci. U.S.A. 81:5575-5579(1984).
RN   [3]
RP   SEQUENCE.
RC   SPECIES=L.decemlineata; TISSUE=Corpora cardiaca;
```

```
RX   MEDLINE=90160053; PubMed=2576128;
RA   Gaede G., Kellner R.;
RT   "The metabolic neuropeptides of the corpus cardiacum from the potato
RT   beetle and the American cockroach are identical.";
RL   Peptides 10:1287-1289(1989).
RN   [4]
RP   SEQUENCE.
RC   SPECIES=B.orientalis; TISSUE=Corpora cardiaca;
RX   MEDLINE=90253659; PubMed=2340112;
RA   Gaede G., Rinehart K.L. Jr.;
RT   "Primary structures of hypertrehalosaemic neuropeptides isolated from
RT   the corpora cardiaca of the cockroaches Leucophaea maderae,
RT   Gromphadorhina portentosa, Blattella germanica and Blatta orientalis
RT   and of the stick insect Extatosoma tiaratum assigned by tandem fast
RT   atom bombardment mass spectrometry.";
RL   Biol. Chem. Hoppe-Seyler 371:345-354(1990).
CC   -!- FUNCTION: HYPERTREHALOSAEMIC FACTORS ARE NEUROPEPTIDES THAT
CC       ELEVATE THE LEVEL OF TREHALOSE IN THE HEMOLYMPH (TREHALOSE IS
CC       THE MAJOR CARBOHYDRATE IN THE HEMOLYMPH OF INSECTS).
CC   -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR   PIR; A05170; A05170.
DR   PIR; S08996; S08996.
DR   PIR; B44960; B44960.
DR   PIR; B49823; B49823.
DR   InterPro; IPR002047; AKH.
DR   PROSITE; PS00256; AKH; 1.
KW   Neuropeptide; Amidation.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  1006 MW;  86745771A9D1A736 CRC64;


  Query Match             21.7%;  Score 13;  DB 1;  Length 8;
  Best Local Similarity   75.0%;  Pred. No. 1e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 LTHT 5
          || |
Db      2 LTFT 5


RESULT  17
PPK1_PERAM
ID   PPK1_PERAM     STANDARD;      PRT;     9 AA.
AC   P82691;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Pyrokinin-1 (Pea-PK-1) (FXPRL-amide).
OS   Periplaneta americana (American cockroach).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC   Blattoidea; Blattidae; Periplaneta.
OX   NCBI_TaxID=6978;
RN   [1]
RP   SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC   TISSUE=Corpora cardiaca;
RX   MEDLINE=97353923; PubMed=9210163;
RA   Predel R., Kellner R., Kaufmann R., Penzlin H., Gaede G.;
RT   "Isolation and structural elucidation of two pyrokinins from the
RT   retrocerebral complex of the American cockroach.";
RL   Peptides 18:473-478(1997).
RN   [2]
RP   TISSUE SPECIFICITY.
RX   MEDLINE=20189894; PubMed=10723010;
RA   Predel R., Eckert M.;
RT   "Tagma-specific distribution of FXPRLamides in the nervous system of
RT   the American cockroach.";
RL   J. Comp. Neurol. 419:352-363(2000).
CC   -!- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
CC       (MYOTROPIC ACTIVITY).
CC   -!- TISSUE SPECIFICITY: CORPORA CARDIACA.
CC   -!- MASS SPECTROMETRY: MW=1010.4; METHOD=MALDI.
CC   -!- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.
DR   InterPro; IPR001484; Pyrokinin.
DR   PROSITE; PS00539; PYROKININ; FALSE_NEG.
KW   Neuropeptide; Amidation; Pyrokinin.
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  1011 MW;  885C176059C87DC1 CRC64;


  Query Match             21.7%;  Score 13;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 HT 5
          ||
Db      1 HT 2


RESULT  18
XYLA_STRSQ
ID   XYLA_STRSQ     STANDARD;      PRT;     9 AA.
AC   P19149;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Xylose isomerase (EC 5.3.1.5) (Fragment).
GN   XYLA.
OS   Streptomyces sp. (strain NCL 82-5-1).
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1931;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=88326335; PubMed=3415697;
RA   Pawar H.S., Kannan K., Srinivasan M.C., Vartak H.G.;
RT   "Purification and characterisation of glucose (xylose) isomerase from
RT   Chainia sp. (NCL 82-5-1).";
RL   Biochem. Biophys. Res. Commun. 155:411-417(1988).
CC   -!- FUNCTION: INVOLVED IN D-XYLOSE CATABOLISM.
CC   -!- CATALYTIC ACTIVITY: D-xylose = D-xylulose.
CC   -!- COFACTOR: Magnesium; binds 2 ions per subunit (Potential).
CC   -!- SUBUNIT: HOMOTETRAMER.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: BELONGS TO THE XYLOSE ISOMERASE FAMILY.
DR   PIR; A31576; A31576.
DR   InterPro; IPR001998; Xylose_isom.
DR   PROSITE; PS00172; XYLOSE_ISOMERASE_1; PARTIAL.
DR   PROSITE; PS00173; XYLOSE_ISOMERASE_2; PARTIAL.
KW   Isomerase; Pentose shunt; Xylose metabolism; Magnesium.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  983 MW;  F64BA1EDC5B87DD1 CRC64;


  Query Match             21.7%;  Score 13;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 HT 5
          ||
Db      7 HT 8


RESULT  19
AMPN_HELAM
ID   AMPN_HELAM     STANDARD;      PRT;    10 AA.
AC   P81731;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Aminopeptidase N (EC 3.4.11.2) (Fragment).
OS   Helicoverpa armigera (Cotton bollworm).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Noctuoidea; Noctuidae; Heliothinae; Helicoverpa.
```

```
OX   NCBI_TaxID=29058;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Midgut;
RA   Ingle S.S., Trivedi N., Prasad R., Rao K.K., Chatpar H.S.;
RT   "Aminopeptidase-N as a receptor for Bacillus thuringiensis CryIAc
RT   toxin from Helicoverpa armigera.";
RL   Submitted (MAR-1999) to the SWISS-PROT data bank.
CC   -!- FUNCTION: ACTS AS A RECEPTOR FOR B.THURIGIENSIS TOXIN CRYIAC.
CC   -!- CATALYTIC ACTIVITY: Release of an N-terminal amino acid,
CC       preferably a neutral or hydrophobic one, from a polypeptide.
CC       Aminoacyl-arylamides are poor substrates.
CC   -!- COFACTOR: BINDS AND REQUIRES A ZINC ATOM, WHICH IS ESSENTIAL FOR
CC       PROTEOLYTIC ACTIVITY (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M1.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   PROSITE; PS00142; ZINC_PROTEASE; PARTIAL.
KW   Hydrolase; Metalloprotease; Aminopeptidase; Zinc.
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1093 MW;  05042EB87B11F1BB CRC64;


  Query Match             21.7%;  Score 13;  DB 1;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+03;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 TH 4
           ||
Db       4 TH 5


RESULT  20
HTF2_CARMO
ID   HTF2_CARMO     STANDARD;      PRT;    10 AA.
AC   P11385;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Hypertrehalosaemic factor II (HTF-II) (HRTH-II) (Hypertrehalosaemic
DE   neuropeptide II).
OS   Carausius morosus (Indian stick insect), and
OS   Extatosoma tiaratum (Stick insect).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Phasmatodea; Heteronemiidae;
OC   Carausius.
OX   NCBI_TaxID=7022, 7024;
RN   [1]
RP   SEQUENCE.
RC   SPECIES=C.morosus; TISSUE=Corpora cardiaca;
RX   MEDLINE=87157103; PubMed=3828078;
RA   Gaede G., Rinehart K.L. Jr.;
RT   "Primary structure of the hypertrehalosaemic factor II from the
RT   corpus cardiacum of the Indian stick insect, Carausius morosus,
RT   determined by fast atom bombardment mass spectrometry.";
RL   Biol. Chem. Hoppe-Seyler 368:67-75(1987).
RN   [2]
RP   SEQUENCE.
RC   SPECIES=E.tiaratum; TISSUE=Corpora cardiaca;
RX   MEDLINE=90253659; PubMed=2340112;
RA   Gaede G., Rinehart K.L. Jr.;
RT   "Primary structures of hypertrehalosaemic neuropeptides isolated from
RT   the corpora cardiaca of the cockroaches Leucophaea maderae,
RT   Gromphadorhina portentosa, Blattella germanica and Blatta orientalis
RT   and of the stick insect Extatosoma tiaratum assigned by tandem fast
RT   atom bombardment mass spectrometry.";
RL   Biol. Chem. Hoppe-Seyler 371:345-354(1990).
RN   [3]
RP   CARBOHYDRATE-LINKAGE SITE.
RC   SPECIES=C.morosus; TISSUE=Corpora cardiaca;
RX   MEDLINE=93129188; PubMed=1482345;
RA   Gaede G., Kellner R., Rinehart K.L. Jr., Proefke M.L.;
RT   "A tryptophan-substituted member of the AKH/RPCH family isolated from
RT   a stick insect corpus cardiacum.";
RL   Biochem. Biophys. Res. Commun. 189:1303-1309(1992).
CC   -!- FUNCTION: HYPERTREHALOSAEMIC FACTORS ARE NEUROPEPTIDES THAT
CC       ELEVATE THE LEVEL OF TREHALOSE IN THE HEMOLYMPH (TREHALOSE IS
CC       THE MAJOR CARBOHYDRATE IN THE HEMOLYMPH OF INSECTS).
CC   -!- MASS SPECTROMETRY: MW=1308.61; METHOD=FAB.
CC   -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR   PIR; S07157; S07157.
DR   PIR; S09138; S09138.
DR   InterPro; IPR002047; AKH.
DR   PROSITE; PS00256; AKH; 1.
KW   Neuropeptide; Amidation; Glycoprotein.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   CARBOHYD      8      8       C-LINKED (MAN) (PROBABLE).
FT   MOD_RES      10     10       AMIDATION.
SQ   SEQUENCE   10 AA;  1164 MW;  9B9036745771A9D1 CRC64;


  Query Match             21.7%;  Score 13;  DB 1;  Length 10;
  Best Local Similarity   75.0%;  Pred. No. 7.8e+03;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 LTHT 5
           || |
Db       2 LTFT 5


Search completed: July  8, 2002, 11:48:58
Job time: 329 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 8, 2002, 11:43:26 ; Search time 21.49 Seconds
(without alignments)
13.639 Million cell updates/sec

Title: US-09-461-061A-3
Perfect score: 56
Sequence: 1 IDNVKKARVQVV 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 83765

Minimum DB seq length: 0
Maximum DB seq length: 12

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	24	42.9	12	1	US-07-995-657-7	Sequence 7, Appli
2	24	42.9	12	1	US-08-474-587-7	Sequence 7, Appli
3	23	41.1	8	4	US-08-160-604-83	Sequence 83, Appl
4	23	41.1	8	4	US-08-160-604-84	Sequence 84, Appl
5	23	41.1	8	4	US-08-160-604-85	Sequence 85, Appl
6	23	41.1	9	2	US-08-908-526-6	Sequence 6, Appli
7	23	41.1	9	3	US-08-159-339A-127	Sequence 127, App
8	23	41.1	10	2	US-08-428-257A-6	Sequence 6, Appli
9	23	41.1	11	1	US-08-378-761A-68	Sequence 68, Appl
10	23	41.1	11	1	US-08-485-286-68	Sequence 68, Appl
11	23	41.1	12	2	US-08-908-526-2	Sequence 2, Appli
12	22	39.3	8	4	US-08-160-604-86	Sequence 86, Appl
13	22	39.3	9	1	US-08-321-668-10	Sequence 10, Appl
14	22	39.3	9	1	US-08-837-941-10	Sequence 10, Appl
15	22	39.3	11	4	US-08-647-405B-6	Sequence 6, Appli
16	22	39.3	12	1	US-08-321-668-15	Sequence 15, Appl
17	22	39.3	12	1	US-08-837-941-15	Sequence 15, Appl
18	21	37.5	9	1	US-08-594-447-42	Sequence 42, Appl
19	21	37.5	9	1	US-08-541-964-41	Sequence 41, Appl
20	21	37.5	9	2	US-08-665-647-56	Sequence 56, Appl
21	21	37.5	10	1	US-08-166-195A-45	Sequence 45, Appl
22	21	37.5	10	2	US-08-436-772-45	Sequence 45, Appl
23	21	37.5	10	2	US-08-436-883B-45	Sequence 45, Appl
24	21	37.5	11	2	US-08-762-695-1	Sequence 1, Appli
25	21	37.5	11	4	US-08-647-405B-7	Sequence 7, Appli
26	21	37.5	12	1	US-08-313-681A-13	Sequence 13, Appl
27	21	37.5	12	1	US-08-321-668-16	Sequence 16, Appl
28	21	37.5	12	1	US-08-321-668-17	Sequence 17, Appl
29	21	37.5	12	1	US-08-321-668-18	Sequence 18, Appl
30	21	37.5	12	1	US-08-321-668-19	Sequence 19, Appl
31	21	37.5	12	1	US-08-837-941-16	Sequence 16, Appl
32	21	37.5	12	1	US-08-837-941-17	Sequence 17, Appl
33	21	37.5	12	1	US-08-837-941-18	Sequence 18, Appl
34	21	37.5	12	1	US-08-837-941-19	Sequence 19, Appl
35	21	37.5	12	3	US-09-322-911-13	Sequence 13, Appl
36	21	37.5	12	4	US-08-687-590-4	Sequence 4, Appli
37	20	35.7	5	4	US-09-187-859-3201	Sequence 3201, Ap
38	20	35.7	6	4	US-09-187-859-3711	Sequence 3711, Ap
39	20	35.7	7	4	US-09-187-859-3712	Sequence 3712, Ap
40	20	35.7	7	4	US-09-187-859-3714	Sequence 3714, Ap
41	20	35.7	8	4	US-08-160-604-82	Sequence 82, Appl
42	20	35.7	8	4	US-09-187-859-3715	Sequence 3715, Ap
43	20	35.7	8	4	US-09-187-859-3717	Sequence 3717, Ap
44	20	35.7	9	2	US-08-704-655-3	Sequence 3, Appli
45	20	35.7	9	3	US-08-159-339A-128	Sequence 128, App

ALIGNMENTS

```
RESULT   1
US-07-995-657-7
; Sequence 7, Application US/07995657
; Patent No. 5443974
;  GENERAL INFORMATION:
;    APPLICANT:  Hitz, William D.
;    APPLICANT:  Yadav, Narendra S.
;    APPLICANT:  Perez_Grau, Luis
;    TITLE OF INVENTION:  Nucleotide Sequence of
;    TITLE OF INVENTION:  Soybean Stearoyl-ACP
;    TITLE OF INVENTION:  Desaturase cDNA
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  E. I. du Pont de Nemours
;      ADDRESSEE:  and Company
;      STREET:  1007 Market Street
;      CITY:  Wilmington
;      STATE:  Delaware
;      COUNTRY:  U.S.A.
;      ZIP:  19898
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.0 MB
;      COMPUTER:  Macintosh
;      OPERATING SYSTEM:  Macintosh System, 6.0
;      SOFTWARE:  Microsoft Word, 4.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/995,657
;      FILING DATE:  19921211
;      FILING DATE:  C)  CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  GEIGER, KATHLEEN W.
;      REGISTRATION NUMBER:  35,880
;      REFERENCE/DOCKET NUMBER:  BB_1022-B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  302-892-8112
;      TELEFAX:  302-892-7949
;      TELEX:  835420
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  12 amino acids
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-995-657-7
```

```
  Query Match             42.9%;  Score 24;  DB 1;  Length 12;
  Best Local Similarity   50.0%;  Pred. No. 1.3e+02;
  Matches    3;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 IDNVKK 6
          ::|:||
Db      3 VENIKK 8


RESULT   2
US-08-474-587-7
; Sequence 7, Application US/08474587
; Patent No. 5760206
;  GENERAL INFORMATION:
;    APPLICANT:  Hitz, William D.
;    APPLICANT:  Yadav, Narendra S.
;    APPLICANT:  Perez_Grau, Luis
;    TITLE OF INVENTION:  Nucleotide Sequence of
; Patent No. 5760206
;    TITLE OF INVENTION:  Soybean Stearoyl-ACP
;    TITLE OF INVENTION:  Desaturase cDNA
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  E. I. du Pont de Nemours
;      ADDRESSEE:  and Company
;      STREET:  1007 Market Street
;      CITY:  Wilmington
;      STATE:  Delaware
;      COUNTRY:  U.S.A.
;      ZIP:  19898
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.0 MB
;      COMPUTER:  Macintosh
;      OPERATING SYSTEM:  Macintosh System, 6.0
;      SOFTWARE:  Microsoft Word, 4.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/474,587
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  SIEGELL, BARBARA C.
;      REGISTRATION NUMBER:  30,684
;      REFERENCE/DOCKET NUMBER:  BB_1022-C
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  302-992-4931
;      TELEFAX:  302-773-0164
;      TELEX:  835420
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  12 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-474-587-7


  Query Match             42.9%;  Score 24;  DB 1;  Length 12;
  Best Local Similarity   50.0%;  Pred. No. 1.3e+02;
  Matches    3;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 IDNVKK 6
          ::|:||
Db      3 VENIKK 8


RESULT   3
US-08-160-604-83
; Sequence 83, Application US/08160604
; Patent No. 6232522
;  GENERAL INFORMATION:
;    APPLICANT:  Harley, John
;    APPLICANT:  James, Judith A.
;    APPLICANT:  Scofield, R. H.
;    TITLE OF INVENTION:  PEPTIDE INDUCTION OF AUTOIMMUNITYAND CLINICAL SYMPTOMATOLOG
;    NUMBER OF SEQUENCES:  127
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patrea L. Pabst
;      STREET:  1100 Peachtree Street, Suite 2800
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30309-4530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/160,604
;      FILING DATE:  30-NOV-1993
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/867,819
;      FILING DATE:  13-APR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/648,205
;      FILING DATE:  31-JAN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/472,947
;      FILING DATE:  31-JAN-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  OMRF114CIP(3)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404)-815-6508
;      TELEFAX:  (404)-815-6555
;  INFORMATION FOR SEQ ID NO:  83:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
US-08-160-604-83


  Query Match             41.1%;  Score 23;  DB 4;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 1.7e+05;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 VKKARV 9
          :||||:
Db      3 LKKARI 8


RESULT   4
US-08-160-604-84
; Sequence 84, Application US/08160604
; Patent No. 6232522
;  GENERAL INFORMATION:
;    APPLICANT:  Harley, John
;    APPLICANT:  James, Judith A.
;    APPLICANT:  Scofield, R. H.
;    TITLE OF INVENTION:  PEPTIDE INDUCTION OF AUTOIMMUNITYAND CLINICAL SYMPTOMATOLOG
;    NUMBER OF SEQUENCES:  127
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patrea L. Pabst
;      STREET:  1100 Peachtree Street, Suite 2800
```

```
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30309-4530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/160,604
;      FILING DATE:  30-NOV-1993
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/867,819
;      FILING DATE:  13-APR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/648,205
;      FILING DATE:  31-JAN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/472,947
;      FILING DATE:  31-JAN-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  OMRF114CIP(3)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404)-815-6508
;      TELEFAX:  (404)-815-6555
;  INFORMATION FOR SEQ ID NO:  84:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
US-08-160-604-84


  Query Match             41.1%;  Score 23;  DB 4;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 1.7e+05;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 VKKARV 9
           :||||:
Db       2 LKKARI 7


RESULT   5
US-08-160-604-85
; Sequence 85, Application US/08160604
; Patent No. 6232522
;  GENERAL INFORMATION:
;    APPLICANT:  Harley, John
;    APPLICANT:  James, Judith A.
;    APPLICANT:  Scofield, R. H.
;    TITLE OF INVENTION:  PEPTIDE INDUCTION OF AUTOIMMUNITYAND CLINICAL SYMPTOMATOLOGY
;    NUMBER OF SEQUENCES:  127
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patrea L. Pabst
;      STREET:  1100 Peachtree Street, Suite 2800
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30309-4530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/160,604
;      FILING DATE:  30-NOV-1993
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/867,819
;      FILING DATE:  13-APR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/648,205
;      FILING DATE:  31-JAN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/472,947
;      FILING DATE:  31-JAN-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  OMRF114CIP(3)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404)-815-6508
;      TELEFAX:  (404)-815-6555
;  INFORMATION FOR SEQ ID NO:  85:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
US-08-160-604-85


  Query Match             41.1%;  Score 23;  DB 4;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 1.7e+05;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 VKKARV 9
           :||||:
Db       1 LKKARI 6


RESULT   6
US-08-908-526-6
; Sequence 6, Application US/08908526
; Patent No. 5876730
;  GENERAL INFORMATION:
;    APPLICANT:  Brigstock, David R.
;    APPLICANT:  Harding, Paul H.
;    TITLE OF INVENTION:  HEPARIN BINDING GROWTH FACTOR (HBGF)
;    TITLE OF INVENTION:  POLYPEPTIDES
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  4225 Executive Square, Suite 1400
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/908,526
;      FILING DATE:  07-AUG-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Haile, Lisa A., Ph.D.
;      REGISTRATION NUMBER:  38,347
```

```
;      REFERENCE/DOCKET NUMBER:  08766/003002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619/678-5070
;      TELEFAX:  619/678-5099
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-908-526-6


  Query Match             41.1%;  Score 23;  DB 2;  Length 9;
  Best Local Similarity   42.9%;  Pred. No. 1.7e+05;
  Matches    3;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 DNVKKAR 8
            :|:|| :
Db       2 ENIKKGK 8


RESULT   7
US-08-159-339A-127
; Sequence 127, Application US/08159339A
; Patent No. 6037135
;  GENERAL INFORMATION:
;    APPLICANT:  Kubo, Ralph T.
;    APPLICANT:  Grey, Howard M.
;    APPLICANT:  Sette, Alessandro
;    APPLICANT:  Celis, Esteban
;    TITLE OF INVENTION:  HLA Binding peptides and Their
;    TITLE OF INVENTION:  Uses
;    NUMBER OF SEQUENCES:  1254
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/159,339A
;      FILING DATE:  29-NOV-1993
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/926,666
;      FILING DATE:  07-AUG-1992
;      APPLICATION NUMBER:  US 08/027,746
;      FILING DATE:  05-MAR-1993
;      APPLICATION NUMBER:  US 08/103,396
;      FILING DATE:  06-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weber, Ellen Lauver
;      REGISTRATION NUMBER:  32,762
;      REFERENCE/DOCKET NUMBER:  018623-005030US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  127:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-159-339A-127


  Query Match             41.1%;  Score 23;  DB 3;  Length 9;
  Best Local Similarity   83.3%;  Pred. No. 1.7e+05;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 IDNVKK 6
            |||| |
Db       3 IDNVYK 8


RESULT   8
US-08-428-257A-6
; Sequence 6, Application US/08428257A
; Patent No. 5885808
;  GENERAL INFORMATION:
;    APPLICANT:  Spooner, Robert A.
;    APPLICANT:  Epenetos, A.A.
;    TITLE OF INVENTION:  Compounds to target cells
;    NUMBER OF SEQUENCES:  80
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jules E. Goldberg
;      STREET:  261 Madison Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10016-2391
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/428,257A
;      FILING DATE:  07/05/95
;      CLASSIFICATION:  514
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-428-257A-6


  Query Match             41.1%;  Score 23;  DB 2;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 1.6e+02;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 VKKARVQV 11
            :|| :||:
Db       2 LKKTKVQL 9


RESULT   9
US-08-378-761A-68
; Sequence 68, Application US/08378761A
; Patent No. 5635384
;  GENERAL INFORMATION:
;    APPLICANT:  WALSH, TERENCE A
;    APPLICANT:  HEY, TIMOTHY D
;    APPLICANT:  MORGAN, ALICE ER
;    TITLE OF INVENTION:  RIBOSOME-INACTIVATING PROTEINS, INACTIVE
;    TITLE OF INVENTION:  PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
;    TITLE OF INVENTION:  USING
;    NUMBER OF SEQUENCES:  81
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ANDREA T. BORUCKI
;      STREET:  9330 ZIONSVILLE ROAD
;      CITY:  INDIANAPOLIS
;      STATE:  IN
;      COUNTRY:  US
;      ZIP:  46268
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/378,761A
;      FILING DATE:  26-JAN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BORUCKI, ANDREA T
;      REGISTRATION NUMBER:  33651
;      REFERENCE/DOCKET NUMBER:  38272B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (317) 337-4846
;  INFORMATION FOR SEQ ID NO:  68:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-378-761A-68


  Query Match             41.1%;  Score 23;  DB 1;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 KKARV 9
           |||||
Db       3 KKARV 7


RESULT  10
US-08-485-286-68
; Sequence 68, Application US/08485286
; Patent No. 5646026
; Patent No. 5646026 5646119
;  GENERAL INFORMATION:
;    APPLICANT:  WALSH, TERENCE A
;    APPLICANT:  HEY, TIMOTHY D
;    APPLICANT:  MORGAN, ALICE ER
;    TITLE OF INVENTION:  RIBOSOME-INACTIVATING PROTEINS, INACTIVE
;    TITLE OF INVENTION:  PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
;    TITLE OF INVENTION:  USING
;    NUMBER OF SEQUENCES:  81
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ANDREA T. BORUCKI
;      STREET:  9330 ZIONSVILLE ROAD
;      CITY:  INDIANAPOLIS
;      STATE:  IN
;      COUNTRY:  US
;      ZIP:  46268
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,286
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/378761
;      FILING DATE:  26-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BORUCKI, ANDREA T
;      REGISTRATION NUMBER:  33651
;      REFERENCE/DOCKET NUMBER:  38272B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (317) 337-4846
;  INFORMATION FOR SEQ ID NO:  68:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-485-286-68


  Query Match             41.1%;  Score 23;  DB 1;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 KKARV 9
           |||||
Db       3 KKARV 7


RESULT  11
US-08-908-526-2
; Sequence 2, Application US/08908526
; Patent No. 5876730
;  GENERAL INFORMATION:
;    APPLICANT:  Brigstock, David R.
;    APPLICANT:  Harding, Paul H.
;    TITLE OF INVENTION:  HEPARIN BINDING GROWTH FACTOR (HBGF)
;    TITLE OF INVENTION:  POLYPEPTIDES
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  4225 Executive Square, Suite 1400
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/908,526
;      FILING DATE:  07-AUG-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Haile, Lisa A., Ph.D.
;      REGISTRATION NUMBER:  38,347
;      REFERENCE/DOCKET NUMBER:  08766/003002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619/678-5070
;      TELEFAX:  619/678-5099
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  12 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-908-526-2


  Query Match             41.1%;  Score 23;  DB 2;  Length 12;
  Best Local Similarity   42.9%;  Pred. No. 1.9e+02;
  Matches    3;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 DNVKKAR 8
             :|:|| :
Db       1 ENIKKGK 7


RESULT  12
US-08-160-604-86
```

```
; Sequence 86, Application US/08160604
; Patent No. 6232522
;  GENERAL INFORMATION:
;    APPLICANT:  Harley, John
;    APPLICANT:  James, Judith A.
;    APPLICANT:  Scofield, R. H.
;    TITLE OF INVENTION:  PEPTIDE INDUCTION OF AUTOIMMUNITYAND CLINICAL SYMPTOMATOLOGY
;    NUMBER OF SEQUENCES:  127
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patrea L. Pabst
;      STREET:  1100 Peachtree Street, Suite 2800
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30309-4530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/160,604
;      FILING DATE:  30-NOV-1993
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/867,819
;      FILING DATE:  13-APR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/648,205
;      FILING DATE:  31-JAN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/472,947
;      FILING DATE:  31-JAN-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  OMRF114CIP(3)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404)-815-6508
;      TELEFAX:  (404)-815-6555
;  INFORMATION FOR SEQ ID NO:  86:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
US-08-160-604-86


  Query Match             39.3%;  Score 22;  DB 4;  Length 8;
  Best Local Similarity   80.0%;  Pred. No. 1.7e+05;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 KKARV 9
            ||||:
Db       1 KKARI 5


RESULT  13
US-08-321-668-10
; Sequence 10, Application US/08321668
; Patent No. 5665859
;  GENERAL INFORMATION:
;    APPLICANT:  WALLACH, David
;    APPLICANT:  BRAKEBUSCH, Cord
;    APPLICANT:  VARFOLOMEEV, Eugene
;    APPLICANT:  BATKIN, Michael
;    TITLE OF INVENTION:  MOLECULES INFLUENCING THE SHEDDING OF
;    TITLE OF INVENTION:  THE TNF RECEPTORS, THEIR PREPARATION AND THEIR USE
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/321,668
;      FILING DATE:  12-OCT-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  IL 107268
;      FILING DATE:  12-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROWDY, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  WALLACH=13
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      OTHER INFORMATION:  /note= hu p55 TNF-R mutant
;      OTHER INFORMATION:  construct  175-179
US-08-321-668-10


  Query Match             39.3%;  Score 22;  DB 1;  Length 9;
  Best Local Similarity   55.6%;  Pred. No. 1.7e+05;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 IDNVKKARV 9
             |:|||   |
Db       1 IENVKGTTV 9


RESULT  14
US-08-837-941-10
; Sequence 10, Application US/08837941
; Patent No. 5766917
;  GENERAL INFORMATION:
;    APPLICANT:  WALLACH, David
;    APPLICANT:  BRAKEBUSCH, Cord
;    APPLICANT:  VARFOLOMEEV, Eugene
;    APPLICANT:  BATKIN, Michael
;    TITLE OF INVENTION:  MOLECULES INFLUENCING THE SHEDDING OF
;    TITLE OF INVENTION:  THE TNF RECEPTORS, THEIR PREPARATION AND THEIR USE
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/837,941
;      FILING DATE:  28-APR-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/321,668
;      FILING DATE:  12-OCT-1994
;      APPLICATION NUMBER:  IL 107268
;      FILING DATE:  12-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROWDY, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  WALLACH=13
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      OTHER INFORMATION:  /note= hu p55 TNF-R mutant
;      OTHER INFORMATION:  construct  175-179
US-08-837-941-10


  Query Match             39.3%;  Score 22;  DB 1;  Length 9;
  Best Local Similarity   55.6%;  Pred. No. 1.7e+05;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 IDNVKKARV 9
            |:|||   |
Db       1 IENVKGTTV 9


RESULT  15
US-08-647-405B-6
; Sequence 6, Application US/08647405B
; Patent No. 6228654
; GENERAL INFORMATION:
;  APPLICANT:  Chait, Brian T.
;  APPLICANT:  Zhao, Yingming
;  APPLICANT:  Kent, Stephen B.H.
;  TITLE OF INVENTION:  METHODS FOR STRUCTURE ANALYSIS OF OLIGOSACCHARIDES
;  FILE REFERENCE:  Oligosaccharides
;  CURRENT APPLICATION NUMBER: US/08/647,405B
;  CURRENT FILING DATE:  1996-05-09
;  NUMBER OF SEQ ID NOS:  7
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 11
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:.Related to
;   OTHER INFORMATION: human translationally controlled tumor protein
US-08-647-405B-6


  Query Match             39.3%;  Score 22;  DB 4;  Length 11;
  Best Local Similarity   57.1%;  Pred. No. 2.7e+02;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 DNVKKAR 8
            | |:||:
Db       5 DTVEKAK 11


RESULT  16
US-08-321-668-15
; Sequence 15, Application US/08321668
; Patent No. 5665859
;  GENERAL INFORMATION:
;    APPLICANT:  WALLACH, David
;    APPLICANT:  BRAKEBUSCH, Cord
;    APPLICANT:  VARFOLOMEEV, Eugene
;    APPLICANT:  BATKIN, Michael
;    TITLE OF INVENTION:  MOLECULES INFLUENCING THE SHEDDING OF
;    TITLE OF INVENTION:  THE TNF RECEPTORS, THEIR PREPARATION AND THEIR USE
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/321,668
;      FILING DATE:  12-OCT-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  IL 107268
;      FILING DATE:  12-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROWDY, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  WALLACH=13
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  12 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      OTHER INFORMATION:  /note= hu p55 TNF-R mutant
;      OTHER INFORMATION:  construct  175-176
US-08-321-668-15


  Query Match             39.3%;  Score 22;  DB 1;  Length 12;
  Best Local Similarity   66.7%;  Pred. No. 3e+02;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 IDNVKK 6
            |:|||:
Db       1 IENVKE 6


RESULT  17
US-08-837-941-15
; Sequence 15, Application US/08837941
; Patent No. 5766917
;  GENERAL INFORMATION:
;    APPLICANT:  WALLACH, David
;    APPLICANT:  BRAKEBUSCH, Cord
;    APPLICANT:  VARFOLOMEEV, Eugene
;    APPLICANT:  BATKIN, Michael
;    TITLE OF INVENTION:  MOLECULES INFLUENCING THE SHEDDING OF
;    TITLE OF INVENTION:  THE TNF RECEPTORS, THEIR PREPARATION AND THEIR USE
;    NUMBER OF SEQUENCES:  42
```

```
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/837,941
;      FILING DATE:  28-APR-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/321,668
;      FILING DATE:  12-OCT-1994
;      APPLICATION NUMBER:  IL 107268
;      FILING DATE:  12-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROWDY, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  WALLACH=13
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  12 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      OTHER INFORMATION:  /note= hu p55 TNF-R mutant
;      OTHER INFORMATION:  construct  175-176
US-08-837-941-15


  Query Match             39.3%;  Score 22;  DB 1;  Length 12;
  Best Local Similarity   66.7%;  Pred. No. 3e+02;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 IDNVKK 6
           |:|||:
Db       1 IENVKE 6


RESULT  18
US-08-594-447-42
; Sequence 42, Application US/08594447
; Patent No. 5776716
;  GENERAL INFORMATION:
;    APPLICANT:  Ron, Dorit
;    APPLICANT:  Napolitano, Eugene W.
;    APPLICANT:  Voronova, Anna F.
;    TITLE OF INVENTION:  METHODS FOR IDENTIFYING AGENTS WHICH
;    TITLE OF INVENTION:  BLOCKTHE INTERACTION OF FYN WITH PKC-THETA, AND USES
;    TITLE OF INVENTION:  THEREOF
;    NUMBER OF SEQUENCES:  75
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Avenue, NW - Ste. 5500
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006-1888
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/594,447
;      FILING DATE:  31-JAN-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murashige, Kate H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  22550-20025.24
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 822-0168
;      TELEX:  90-4030 MRSNFOERSWSH
;  INFORMATION FOR SEQ ID NO:  42:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Peptide
;      LOCATION:  1..9
;      OTHER INFORMATION:  /label= epsilon-V3-3
US-08-594-447-42


  Query Match             37.5%;  Score 21;  DB 1;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 1.7e+05;
  Matches    3;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 DNVKKA 7
           :|:::||
Db       2 NNIRKA 7


RESULT  19
US-08-541-964-41
; Sequence 41, Application US/08541964
; Patent No. 5783405
;  GENERAL INFORMATION:
;    APPLICANT:  Mochly-Rosen, Daria
;    APPLICANT:  Ron, Dorit
;    APPLICANT:  Kauvar, Lawrence M.
;    APPLICANT:  Napolitano, Eugene W.
;    TITLE OF INVENTION:  A RAPID SCREENING METHOD FOR EFFECTORS
;    TITLE OF INVENTION:  OF SIGNAL TRANSDUCTION
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 PENNSYLVANIA AVENUE, NW-STE. 5500
;      CITY:  WASHINGTON
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006-1888
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/541,964
;      FILING DATE:  10-OCT-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murashige, Kate H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  22550-20025.23
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 822-0168
;      TELEX:  90-4030 MRSNFOERSWSH
```

```
;  INFORMATION FOR SEQ ID NO:  41:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Peptide
;      LOCATION:  1..9
;      OTHER INFORMATION:  /label= epsilon-V3-3
US-08-541-964-41


  Query Match             37.5%;  Score 21;  DB 1;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 1.7e+05;
  Matches    3;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 DNVKKA 7
           :|::||
Db       2 NNIRKA 7


RESULT  20
US-08-665-647-56
; Sequence 56, Application US/08665647
; Patent No. 5935803
;  GENERAL INFORMATION:
;    APPLICANT:  Dasquez, Nicki J.
;    APPLICANT:  Ron, Dorit
;    APPLICANT:  Voronova, Anna F.
;    APPLICANT:  Napolitano, Eugene W.
;    TITLE OF INVENTION:  METHODS TO IDENTIFY IMMUNOMODULATORS
;    TITLE OF INVENTION:  USING COGNATE INTERACTION OF PKC-THETA
;    NUMBER OF SEQUENCES:  89
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Avenue, NW - Ste. 5500
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006-1888
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/665,647
;      FILING DATE:  18-JUN-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murashige, Kate H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  22550-20025.25
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 822-0168
;      TELEX:  90-4030 MRSNFOERSWSH
;  INFORMATION FOR SEQ ID NO:  56:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Peptide
;      LOCATION:  1..9
;      OTHER INFORMATION:  /label= epsilon-V3-3
US-08-665-647-56
```

```
  Query Match             37.5%;  Score 21;  DB 2;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 1.7e+05;
  Matches    3;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 DNVKKA 7
           :|::||
Db       2 NNIRKA 7


Search completed: July  8, 2002, 11:43:26
Job time: 148 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 8, 2002, 11:38:48 ; Search time 10.15 Seconds
(without alignments)
30.518 Million cell updates/sec

Title: US-09-461-061A-1
Perfect score: 45
Sequence: 1 NNATFYFK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 148

Minimum DB seq length: 0
Maximum DB seq length: 8

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	18	40.0	8	1	UPAA_HUMAN	P30096 homo sapien
2	16	35.6	8	1	LCK8_LEUMA	P19990 leucophaea
3	15	33.3	5	1	PAP2_PARMA	P81864 pardachirus
4	14	31.1	4	1	FFKA_ANTEL	P58705 anthopleura
5	14	31.1	8	1	LCK4_LEUMA	P21143 leucophaea
6	13	28.9	4	1	FYRI_ANTEL	P58706 anthopleura
7	12	26.7	5	1	TPIS_CANFA	P54714 canis famil
8	12	26.7	8	1	NPB_BOVIN	P15507 bos taurus
9	11	24.4	6	1	FARP_MONEX	P41966 moniezia ex
10	11	24.4	7	1	FAR1_PROCL	P38499 procambarus
11	11	24.4	8	1	AKH_TABAT	P14595 tabanus atr
12	11	24.4	8	1	COXG_RAT	P80430 rattus norv
13	11	24.4	8	1	FAR4_HOMAM	P41487 homarus ame
14	11	24.4	8	1	HTF2_PERAM	P04549 periplaneta
15	11	24.4	8	1	RS1_ERWCH	P37985 erwinia chr
16	11	24.4	8	1	UC26_MAIZE	P80632 zea mays (m
17	10	22.2	5	1	UC22_MAIZE	P80628 zea mays (m
18	10	22.2	6	1	TMOF_SARBU	P41495 sarcophaga
19	10	22.2	6	1	UN06_CLOPA	P81351 clostridium
20	10	22.2	7	1	CCF1_ENTFA	P20104 enterococcu
21	10	22.2	7	1	UN06_PINPS	P81675 pinus pinas
22	10	22.2	8	1	LCK6_LEUMA	P19988 leucophaea
23	9	20.0	4	1	FAR3_HIRME	P42562 hirudo medi
24	9	20.0	4	1	FAR4_HIRME	P42563 hirudo medi
25	9	20.0	5	1	FARP_ARTTR	P41853 artioposthi
26	9	20.0	5	1	RE11_LITRU	P82070 litoria rub
27	9	20.0	5	1	RE21_LITRU	P82071 litoria rub
28	9	20.0	5	1	RE31_LITRU	P82072 litoria rub
29	9	20.0	5	1	RE32_LITRU	P82073 litoria rub
30	9	20.0	5	1	SUGA_ACHDO	P19991 acheta dome
31	9	20.0	7	1	CIA_ENTFA	P11932 enterococcu
32	9	20.0	7	1	FAR2_ASCSU	P31890 ascaris suu
33	9	20.0	7	1	FAR3_HAECO	P81298 haemonchus
34	9	20.0	7	1	FAR3_PANRE	P41874 panagrellus
35	9	20.0	7	1	FAR5_HIRME	P42564 hirudo medi
36	9	20.0	8	1	AKHG_GRYBI	P14086 gryllus bim
37	9	20.0	8	1	AKH_LIBAU	P25418 libellula a
38	9	20.0	8	1	B44K_PORGI	P81886 porphyromon
39	8	17.8	6	1	OVM_LEPDE	P42985 leptinotars
40	8	17.8	7	1	ALL5_CARMA	P81808 carcinus ma
41	8	17.8	7	1	CHOX_ALCSP	P16101 alcaligenes
42	8	17.8	7	1	HY7_PIG	P01153 sus scrofa
43	8	17.8	7	1	TY51_LITRU	P82065 litoria rub
44	8	17.8	8	1	ALL7_CARMA	P81809 carcinus ma
45	8	17.8	8	1	LPMS_STAEP	P23211 staphylococ

ALIGNMENTS

RESULT 1
UPAA_HUMAN
ID UPAA_HUMAN STANDARD; PRT; 8 AA.
AC P30096;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Unknown protein from 2D-page of plasma (Spot 36) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RC TISSUE=Plasma;
RX MEDLINE=93092937; PubMed=1459097;
RA Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,
RA Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,
RA Hochstrasser D.F.;
RT "Plasma protein map: an update by microsequencing.";
RL Electrophoresis 13:707-714(1992).
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 7, ITS MW IS: 12 kDa.
DR SWISS-2DPAGE; P30096; HUMAN.
FT NON_TER 1 1
FT VARIANT 5 5 F -> P.
FT /FTId=VAR_000004.
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 909 MW; 86677B59D1A72042 CRC64;

Query Match 40.0%; Score 18; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TFY 6
|||
Db 4 TFY 6

RESULT 2
LCK8_LEUMA
ID LCK8_LEUMA STANDARD; PRT; 8 AA.
AC P19990;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Leucokinin VIII (L-VIII).
OS Leucophaea maderae (Madeira cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberoidea; Blaberidae; Leucophaea.
OX NCBI_TaxID=6988;
RN [1]
RP SEQUENCE.
RC TISSUE=Head;

```
RA   Holman G.M., Cook B.J., Nachman R.J.;
RT   "Isolation, primary structure and synthesis of leucokinins VII and
RT   VIII: the final members of this new family of cephalomyotropic
RT   peptides isolated from head extracts of Leucophaea maderae.";
RL   Comp. Biochem. Physiol. 88C:31-34(1987).
CC   -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
CC       ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).
CC   -!- SIMILARITY: TO THE OTHER LEUCOKININS.
DR   PIR; JS0318; JS0318.
KW   Neuropeptide; Amidation.
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  902 MW;  736365AB59CAADD8 CRC64;


  Query Match             35.6%;  Score 16;  DB 1;  Length 8;
  Best Local Similarity   75.0%;  Pred. No. 1e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 ATFY 6
          | ||
Db      2 ADFY 5


RESULT   3
PAP2_PARMA
ID   PAP2_PARMA     STANDARD;      PRT;     5 AA.
AC   P81864;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Pardaxin II (PXII) (Fragment).
OS   Pardachirus marmoratus (Red sea moses sole).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC   Soleoidei; Soleidae; Pardachirus.
OX   NCBI_TaxID=31087;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Skin secretion;
RX   MEDLINE=87057369; PubMed=3782138;
RA   Lazarovici P., Primor N., Loew L.M.;
RT   "Purification and pore-forming activity of two hydrophobic
RT   polypeptides from the secretion of the Red sea moses sole (Pardachirus
RT   marmoratus).";
RL   J. Biol. Chem. 261:16704-16713(1986).
CC   -!- FUNCTION: EXHIBITS UNUSUAL SHARK REPELLENT AND SURFACTANT
CC       PROPERTIES. FORMS VOLTAGE-DEPENDENT, ION-PERMEABLE CHANNELS
CC       IN MEMBRANES. AT HIGH CONCENTRATION CAUSES CELL MEMBRANE LYSIS.
CC   -!- SUBUNIT: MONOMER. IN AQUEOUS SOLUTION EXISTS AS A TETRAMER.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE PARDAXIN FAMILY.
KW   Toxin.
FT   NON_TER       5      5
SQ   SEQUENCE   5 AA;  614 MW;  7769C9C9C8100000 CRC64;


  Query Match             33.3%;  Score 15;  DB 1;  Length 5;
  Best Local Similarity   66.7%;  Pred. No. 1e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 FYF 7
          |:|
Db      2 FFF 4


RESULT   4
FFKA_ANTEL
ID   FFKA_ANTEL     STANDARD;      PRT;     4 AA.
AC   P58705;
DT   01-MAR-2002 (Rel. 41, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Antho-KAamide.
OS   Anthopleura elegantissima (Sea anemone).
OC   Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
OC   Nynantheae; Actiniidae; Anthopleura.
OX   NCBI_TaxID=6110;
RN   [1]
RP   SEQUENCE.
RX   PubMed=1681803;
RA   Nothacker H.-P., Rinehart K.L. Jr., Grimmelikhuijzen C.J.P.;
RT   "Isolation of L-3-phenyllactyl-Phe-Lys-Ala-NH2 (Antho-KAamide), a
RT   novel neuropeptide from sea anemones.";
RL   Biochem. Biophys. Res. Commun. 179:1205-1211(1991).
RN   [2]
RP   FUNCTION.
RX   PubMed=8397415;
RA   McFarlane I.D., Hudman D., Nothacker H.-P., Grimmelikhuijzen C.J.P.;
RT   "The expansion behaviour of sea anemones may be coordinated by two
RT   inhibitory neuropeptides, Antho-KAamide and Antho-RIamide.";
RL   Proc. R. Soc. Lond., B, Biol. Sci. 253:183-188(1993).
CC   -!- FUNCTION: Inhibits spontaneous contractions in several muscle
CC       groups. May be involved in the expansion phase of feeding
CC       behaviour in sea anemones.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Neuron-specific.
KW   Neuropeptide; Amidation.
FT   MOD_RES       1      1       L-3-PHENYLLACTYL.
FT   MOD_RES       4      4       AMIDATION.
SQ   SEQUENCE   4 AA;  512 MW;  6DD339C9A0000000 CRC64;


  Query Match             31.1%;  Score 14;  DB 1;  Length 4;
  Best Local Similarity   66.7%;  Pred. No. 1e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      6 YFK 8
          :||
Db      1 FFK 3


RESULT   5
LCK4_LEUMA
ID   LCK4_LEUMA     STANDARD;      PRT;     8 AA.
AC   P21143;
DT   01-MAY-1991 (Rel. 18, Created)
DT   01-MAY-1991 (Rel. 18, Last sequence update)
DT   01-MAY-1991 (Rel. 18, Last annotation update)
DE   Leucokinin IV (L-IV).
OS   Leucophaea maderae (Madeira cockroach).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC   Blaberoidea; Blaberidae; Leucophaea.
OX   NCBI_TaxID=6988;
RN   [1]
RP   SEQUENCE, AND SYNTHESIS.
RC   TISSUE=Head;
RA   Holman G.M., Cook B.J., Nachman R.J.;
RT   "Primary structure and synthesis of two additional neuropeptides
RT   from Leucophaea maderae: members of a new family of
RT   Cephalomyotropins.";
RL   Comp. Biochem. Physiol. 84C:271-276(1986).
CC   -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
CC       ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).
CC   -!- SIMILARITY: TO THE OTHER LEUCOKININS.
KW   Neuropeptide; Amidation.
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  906 MW;  DC6365B1E9D5BDDA CRC64;


  Query Match             31.1%;  Score 14;  DB 1;  Length 8;
  Best Local Similarity   40.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 NATFY 6
```

```
          :|:|:
Db      1 DASFH 5


RESULT   6
FYRI_ANTEL
ID   FYRI_ANTEL     STANDARD;      PRT;     4 AA.
AC   P58706;
DT   01-MAR-2002 (Rel. 41, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Antho-RIamide I [Contains: Antho-RIamide II].
OS   Anthopleura elegantissima (Sea anemone).
OC   Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
OC   Nynantheae; Actiniidae; Anthopleura.
OX   NCBI_TaxID=6110;
RN   [1]
RP   SEQUENCE.
RX   PubMed=1821096;
RA   Nothacker H.-P., Rinehart K.L. Jr., McFarlane I.D.,
RA   Grimmelikhuijzen C.J.P.;
RT   "Isolation of two novel neuropeptides from sea anemones: the unusual,
RT   biologically active L-3-phenyllactyl-Tyr-Arg-Ile-NH2 and its
RT   des-phenyllactyl fragment Tyr-Arg-Ile-NH2.";
RL   Peptides 12:1165-1173(1991).
RN   [2]
RP   FUNCTION.
RX   PubMed=8397415;
RA   McFarlane I.D., Hudman D., Nothacker H.-P., Grimmelikhuijzen C.J.P.;
RT   "The expansion behaviour of sea anemones may be coordinated by two
RT   inhibitory neuropeptides, Antho-KAamide and Antho-RIamide.";
RL   Proc. R. Soc. Lond., B, Biol. Sci. 253:183-188(1993).
CC   -!- FUNCTION: Inhibits spontaneous contractions in several muscle
CC       groups. May be involved in the expansion phase of feeding
CC       behaviour in sea anemones.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Neuron-specific.
KW   Neuropeptide; Amidation.
FT   CHAIN         1      4       ANTHO-RIAMIDE I.
FT   CHAIN         2      4       ANTHO-RIAMIDE II.
FT   MOD_RES       1      1       L-3-PHENYLLACTYL.
FT   MOD_RES       4      4       AMIDATION.
SQ   SEQUENCE   4 AA;  598 MW;  60441B59A0000000 CRC64;


  Query Match             28.9%;  Score 13;  DB 1;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 FY 6
          ||
Db      1 FY 2


RESULT   7
TPIS_CANFA
ID   TPIS_CANFA     STANDARD;      PRT;     5 AA.
AC   P54714;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Triosephosphate isomerase (EC 5.3.1.1) (TIM) (Fragment).
GN   TPI1.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Heart;
RX   MEDLINE=98163340; PubMed=9504812;
RA   Dunn M.J., Corbett J.M., Wheeler C.H.;
RT   "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
RT   dog heart proteins.";
RL   Electrophoresis 18:2795-2802(1997).
CC   -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate = glycerone
CC       phosphate.
CC   -!- PATHWAY: PLAYS AN IMPORTANT ROLE IN SEVERAL METABOLIC PATHWAYS.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE TRIOSEPHOSPHATE ISOMERASE FAMILY.
DR   HSC-2DPAGE; P54714; DOG.
DR   InterPro; IPR000652; Trioseph_isomerase.
DR   PROSITE; PS00171; TIM; PARTIAL.
KW   Isomerase; Glycolysis; Gluconeogenesis; Fatty acid biosynthesis;
KW   Pentose shunt.
FT   NON_TER       1      1
FT   NON_TER       5      5
SQ   SEQUENCE   5 AA;  550 MW;  64444862C9A00000 CRC64;


  Query Match             26.7%;  Score 12;  DB 1;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 NN 2
          ||
Db      4 NN 5


RESULT   8
NPB_BOVIN
ID   NPB_BOVIN      STANDARD;      PRT;     8 AA.
AC   P15507;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-APR-1990 (Rel. 14, Last annotation update)
DE   Neuropeptide B.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Brain;
RX   MEDLINE=86067985; PubMed=3865193;
RA   Yang H.-Y.T., Fratta W., Majane E.A., Costa E.;
RT   "Isolation, sequencing, synthesis, and pharmacological
RT   characterization of two brain neuropeptides that modulate the action
RT   of morphine.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:7757-7761(1985).
CC   -!- FUNCTION: MODULATES THE ACTION OF MORPHINE.
DR   PIR; B24749; B24749.
KW   Neuropeptide; Amidation.
FT   SIMILAR       5      8       TO NEUROPEPTIDE A (AA 5-8) (IDENTICAL).
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  1082 MW;  87D416C776D9C729 CRC64;


  Query Match             26.7%;  Score 12;  DB 1;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      5 FYFK 8
          | |:
Db      1 FLFQ 4


RESULT   9
FARP_MONEX
ID   FARP_MONEX     STANDARD;      PRT;     6 AA.
AC   P41966;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   FMRFamide-like neuropeptide GNFFRF-amide.
```

```
OS   Moniezia expansa (Sheep tapeworm).
OC   Eukaryota; Metazoa; Platyhelminthes; Turbellarian Platyhelminths;
OC   Rhabditophora; Eulecithophora; Revertospermata; Mediofusata;
OC   Neodermata; Cestoda; Eucestoda; Cyclophyllidea; Anoplocephalidae;
OC   Moniezia.
OX   NCBI_TaxID=28841;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=93312289; PubMed=8323531;
RA   Maule A.G., Shaw C., Halton D.W., Thim L.;
RT   "GNFFRFamide: a novel FMRFamide-immunoreactive peptide isolated from
RT   the sheep tapeworm, Moniezia expansa.";
RL   Biochem. Biophys. Res. Commun. 193:1054-1060(1993).
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES       6       6       AMIDATION.
SQ   SEQUENCE   6 AA;  787 MW;  69D409C9C4481000 CRC64;


  Query Match             24.4%;  Score 11;  DB 1;  Length 6;
  Best Local Similarity   33.3%;  Pred. No. 1e+05;
  Matches    1;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 YFK 8
           :|:
Db       3 FFR 5


RESULT  10
FAR1_PROCL
ID   FAR1_PROCL     STANDARD;      PRT;     7 AA.
AC   P38499;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Cardioexcitatory FMRFamide homolog NF1.
OS   Procambarus clarkii (Red swamp crayfish).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
OC   Astacoidea; Cambaridae; Procambarus.
OX   NCBI_TaxID=6728;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Pericardial organs;
RX   MEDLINE=93248032; PubMed=8387183;
RA   Mercier A.J., Orchard I., Tebrugge V., Skerrett M.;
RT   "Isolation of two FMRFamide-related peptides from crayfish
RT   pericardial organs.";
RL   Peptides 14:137-143(1993).
CC   -!- FUNCTION: INCREASES THE RATE AND AMPLITUDE OF SPONTANEOUS
CC       CONTRACTIONS OF SEMI-ISOLATED HEARTS. INCREASES THE AMPLITUDE OF
CC       EXCITATORY POSTSYNAPTIC POTENTIALS IN ABDOMINAL EXTENSOR MUSCLE.
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES       7       7       AMIDATION.
SQ   SEQUENCE   7 AA;  966 MW;  69D40729C4540420 CRC64;


  Query Match             24.4%;  Score 11;  DB 1;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 NATF 5
           | . |
Db       1 NRNF 4


RESULT  11
AKH_TABAT
ID   AKH_TABAT      STANDARD;      PRT;     8 AA.
AC   P14595;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   Adipokinetic hormone (AKH) (Dipteran corpora cardiaca factor I)
DE   (DCC I).
OS   Tabanus atratus (Horse fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Tabanomorpha;
OC   Tabanidae; Tabanus.
OX   NCBI_TaxID=7207;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Corpora cardiaca;
RX   MEDLINE=90046758; PubMed=2813385;
RA   Jaffe H., Raina A.K., Riley C.T., Fraser B.A., Nachman R.J.,
RA   Vogel V.W., Zhang Y.-S., Hayes D.K.;
RT   "Primary structure of two neuropeptide hormones with adipokinetic and
RT   hypotrehalosemic activity isolated from the corpora cardiaca of horse
RT   flies (Diptera).";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:8161-8164(1989).
CC   -!- FUNCTION: THIS HORMONE, RELEASED FROM CELLS IN THE CORPORA
CC       CARDIACA AFTER THE BEGINNING OF FLIGHT, CAUSES RELEASE OF
CC       DIGLYCERIDES FROM THE FAT BODY AND THEN STIMULATES THE FLIGHT
CC       MUSCLES TO USE THESE DIGLYCERIDES AS AN ENERGY SOURCE.
CC   -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR   PIR; A33995; A33995.
DR   InterPro; IPR002047; AKH.
DR   PROSITE; PS00256; AKH; 1.
KW   Neuropeptide; Amidation; Flight.
FT   MOD_RES       1       1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES       8       8       AMIDATION.
SQ   SEQUENCE   8 AA;  949 MW;  86786771A9D1A736 CRC64;


  Query Match             24.4%;  Score 11;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 TF 5
           ||
Db       3 TF 4


RESULT  12
COXG_RAT
ID   COXG_RAT       STANDARD;      PRT;     8 AA.
AC   P80430;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   Cytochrome c oxidase polypeptide VIb (EC 1.9.3.1) (AED) (Fragment).
GN   COX6B.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=WISTAR; TISSUE=Liver;
RX   MEDLINE=95324529; PubMed=7601105;
RA   Schaegger H., Noack H., Halangk W., Brandt U., von Jagow G.;
RT   "Cytochrome-c oxidase in developing rat heart. Enzymic properties and
RT   amino-terminal sequences suggest identity of the fetal heart and the
RT   adult liver isoform.";
RL   Eur. J. Biochem. 230:235-241(1995).
CC   -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC       CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC       MITOCHONDRIAL ELECTRON TRANSPORT.
CC   -!- FUNCTION: THIS PROTEIN MAY BE ONE OF THE HEME-BINDING SUBUNITS OF
CC       THE OXIDASE.
CC   -!- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC       c + 2 H(2)O.
KW   Oxidoreductase; Mitochondrion.
```

```
FT   NON_TER       1      1
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  1039 MW;  8101E9CAA73AE456 CRC64;


  Query Match             24.4%;  Score 11;  DB 1;  Length 8;
  Best Local Similarity   33.3%;  Pred. No. 1e+05;
  Matches    2;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 NNATFY 6
          |   |:
Db      2 NXLDFH 7


RESULT  13
FAR4_HOMAM
ID   FAR4_HOMAM     STANDARD;      PRT;     8 AA.
AC   P41487;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   FMRFamide-like neuropeptide 4 (FLI 4) (F1).
OS   Homarus americanus (American lobster).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
OC   Nephropoidea; Nephropidae; Homarus.
OX   NCBI_TaxID=6706;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Pericardial organs;
RX   MEDLINE=88116164; PubMed=3429714;
RA   Trimmer B.A., Kobierski L.A., Kravitz E.A.;
RT   "Purification and characterization of FMRFamidelike immunoreactive
RT   substances from the lobster nervous system: isolation and sequence
RT   analysis of two closely related peptides.";
RL   J. Comp. Neurol. 266:16-26(1987).
CC   -!- FUNCTION: CAN ACT AS A MODULATOR OF EXOSKELETAL AND CARDIAC
CC       NEUROMUSCULAR JUNCTIONS.
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  1067 MW;  DDD40729C4540451 CRC64;


  Query Match             24.4%;  Score 11;  DB 1;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 NATF 5
          |  |
Db      2 NRNF 5


RESULT  14
HTF2_PERAM
ID   HTF2_PERAM     STANDARD;      PRT;     8 AA.
AC   P04549;
DT   13-AUG-1987 (Rel. 05, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   Hypertrehalosaemic factor II (Neuropeptide M-II) (Periplanetin CC-2)
DE   (PeA-CAH-II) (LeD-CC-II) (Hypertrehalosaemic neuropeptide II).
OS   Periplaneta americana (American cockroach),
OS   Leptinotarsa decemlineata (Colorado potato beetle), and
OS   Blatta orientalis (Oriental cockroach).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC   Blattoidea; Blattidae; Periplaneta.
OX   NCBI_TaxID=6978, 7539, 6976;
RN   [1]
RP   SEQUENCE.
RC   SPECIES=P.americana;
RX   MEDLINE=85046530; PubMed=6548628;
RA   Witten J.L., Schaffer M.H., O'Shea M., Cook J.C., Hemling M.E.,
RA   Rinehart K.L. Jr.;
RT   "Structures of two cockroach neuropeptides assigned by fast atom
RT   bombardment mass spectrometry.";
RL   Biochem. Biophys. Res. Commun. 124:350-358(1984).
RN   [2]
RP   SEQUENCE.
RC   SPECIES=P.americana;
RX   MEDLINE=84298179; PubMed=6591205;
RA   Scarborough R.M., Jamieson G.C., Kalish F., Kramer S.J., McEnroe G.A.,
RA   Miller C.A., Schooley D.A.;
RT   "Isolation and primary structure of two peptides with
RT   cardioacceleratory and hyperglycemic activity from the corpora
RT   cardiaca of Periplaneta americana.";
RL   Proc. Natl. Acad. Sci. U.S.A. 81:5575-5579(1984).
RN   [3]
RP   SEQUENCE.
RC   SPECIES=L.decemlineata; TISSUE=Corpora cardiaca;
RX   MEDLINE=90160053; PubMed=2576128;
RA   Gaede G., Kellner R.;
RT   "The metabolic neuropeptides of the corpus cardiacum from the potato
RT   beetle and the American cockroach are identical.";
RL   Peptides 10:1287-1289(1989).
RN   [4]
RP   SEQUENCE.
RC   SPECIES=B.orientalis; TISSUE=Corpora cardiaca;
RX   MEDLINE=90253659; PubMed=2340112;
RA   Gaede G., Rinehart K.L. Jr.;
RT   "Primary structures of hypertrehalosaemic neuropeptides isolated from
RT   the corpora cardiaca of the cockroaches Leucophaea maderae,
RT   Gromphadorhina portentosa, Blattella germanica and Blatta orientalis
RT   and of the stick insect Extatosoma tiaratum assigned by tandem fast
RT   atom bombardment mass spectrometry.";
RL   Biol. Chem. Hoppe-Seyler 371:345-354(1990).
CC   -!- FUNCTION: HYPERTREHALOSAEMIC FACTORS ARE NEUROPEPTIDES THAT
CC       ELEVATE THE LEVEL OF TREHALOSE IN THE HEMOLYMPH (TREHALOSE IS
CC       THE MAJOR CARBOHYDRATE IN THE HEMOLYMPH OF INSECTS).
CC   -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR   PIR; A05170; A05170.
DR   PIR; S08996; S08996.
DR   PIR; B44960; B44960.
DR   PIR; B49823; B49823.
DR   InterPro; IPR002047; AKH.
DR   PROSITE; PS00256; AKH; 1.
KW   Neuropeptide; Amidation.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  1006 MW;  86745771A9D1A736 CRC64;


  Query Match             24.4%;  Score 11;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 TF 5
          ||
Db      3 TF 4


RESULT  15
RS1_ERWCH
ID   RS1_ERWCH      STANDARD;      PRT;     8 AA.
AC   P37985;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   30S ribosomal protein S1 (Fragment).
GN   RPSA.
OS   Erwinia chrysanthemi.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Pectobacterium.
OX   NCBI_TaxID=556;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=3937;
RA   Douillie A., Toussaint A., Faelen M.;
RL   Submitted (AUG-1993) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: BINDS MRNA; THUS FACILITATING RECOGNITION OF THE
CC       INITIATION POINT. IT IS NEEDED TO TRANSLATE MRNA WITH A SHORT
CC       SHINE-DALGARNO (SD) PURINE-RICH SEQUENCE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE S1P FAMILY OF RIBOSOMAL PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X74750; CAA52769.1; -.
DR   PIR; S37141; S37141.
KW   Ribosomal protein; Repeat; RNA-binding.
FT   NON_TER       1      1
SQ   SEQUENCE   8 AA;  837 MW;  9E18733DC5B339CD CRC64;


  Query Match             24.4%;  Score 11;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      7 FK 8
          ||
Db      2 FK 3


RESULT  16
UC26_MAIZE
ID   UC26_MAIZE     STANDARD;      PRT;     8 AA.
AC   P80632;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Unknown protein from 2D-page of etiolated coleoptile (Spot 907)
DE   (Fragment).
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC   Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Coleoptile;
RA   Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA   Pernollet J.-C., Zivy M., de Vienne D.;
RT   "The maize two dimensional gel protein database: towards an integrated
RT   genome analysis program.";
RL   Theor. Appl. Genet. 93:997-1005(1996).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 7.0, ITS MW IS: 57.2 kDa.
DR   Maize-2DPAGE; P80632; COLEOPTILE.
FT   NON_TER       1      1
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  990 MW;  9639D6DAB4176B1D CRC64;


  Query Match             24.4%;  Score 11;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      7 FK 8
          ||
Db      7 FK 8


RESULT  17
UC22_MAIZE
ID   UC22_MAIZE     STANDARD;      PRT;     5 AA.
AC   P80628;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Unknown protein from 2D-page of etiolated coleoptile (Spot 474)
DE   (Fragment).
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC   Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Coleoptile;
RA   Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA   Pernollet J.-C., Zivy M., de Vienne D.;
RT   "The maize two dimensional gel protein database: towards an integrated
RT   genome analysis program.";
RL   Theor. Appl. Genet. 93:997-1005(1996).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 6.1, ITS MW IS: 30.4 kDa.
DR   Maize-2DPAGE; P80628; COLEOPTILE.
DR   MaizeDB; 123954; -.
FT   NON_TER       1      1
FT   NON_TER       5      5
SQ   SEQUENCE   5 AA;  654 MW;  72CB19C9C0300000 CRC64;


  Query Match             22.2%;  Score 10;  DB 1;  Length 5;
  Best Local Similarity   33.3%;  Pred. No. 1e+05;
  Matches    1;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      6 YFK 8
          :|:
Db      2 FFE 4


RESULT  18
TMOF_SARBU
ID   TMOF_SARBU     STANDARD;      PRT;     6 AA.
AC   P41495;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   Trypsin-modulating oostatic factor (TMOF).
OS   Sarcophaga bullata (Grey flesh fly) (Neobellieria bullata).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Oestroidea; Sarcophagidae; Sarcophaga.
OX   NCBI_TaxID=7385;
RN   [1]
RP   SEQUENCE, AND SYNTHESIS.
RC   TISSUE=Ovary;
RX   MEDLINE=94211930; PubMed=8159807;
RA   Bylemans D., Borovsky D., Hunt D.F., Shabanowitz J., Grauwels L.,
RA   de Loof A.;
RT   "Sequencing and characterization of trypsin modulating oostatic
RT   factor (TMOF) from the ovaries of the grey fleshfly, Neobellieria
RT   (Sarcophaga) bullata.";
RL   Regul. Pept. 50:61-72(1994).
CC   -!- FUNCTION: HAS AN OOSTATIC ACTIVITY. INHIBITS TRYPSIN BIOSYNTHESIS
CC       IN THE MIDGUT WHICH INDIRECTLY REDUCES THE VITELLOGENIN
CC       CONCENTRATION IN THE HEMOLYMPH RESULTING IN INHIBITION OF OOCYTE
CC       DEVELOPMENT.
CC   -!- DEVELOPMENTAL STAGE: SYNTHESIZED AND RELEASED FROM FOLLICULAR
CC       EPITHELIUM AFTER A BLOOD MEAL.
KW   Hormone.
SQ   SEQUENCE   6 AA;  695 MW;  61E72451B7642000 CRC64;
```

Query Match 22.2%; Score 10; DB 1; Length 6;
Best Local Similarity 66.7%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      2 NAT 4
          | |
Db      1 NPT 3
```

RESULT 19
UN06_CLOPA
ID UN06_CLOPA STANDARD; PRT; 6 AA.
AC P81351;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Unknown protein CP 6 from 2D-page (Fragment).
OS Clostridium pasteurianum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1501;
RN [1]
RP SEQUENCE.
RC STRAIN=W5;
RX MEDLINE=98291870; PubMed=9629918;
RA Flengsrud R., Skjeldal L.;
RT "Two-dimensional gel electrophoresis separation and N-terminal
RT sequence analysis of proteins from Clostridium pasteurianum W5.";
RL Electrophoresis 19:802-806(1998).
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 5.0, ITS MW IS: 75.9 kDa.
FT NON_TER 6 6
SQ SEQUENCE 6 AA; 657 MW; 605B1DC1A45A8000 CRC64;

Query Match 22.2%; Score 10; DB 1; Length 6;
Best Local Similarity 66.7%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 NNA 3
          | |
Db      2 NTA 4
```

RESULT 20
CCF1_ENTFA
ID CCF1_ENTFA STANDARD; PRT; 7 AA.
AC P20104;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Sex pheromone CCF10.
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;
OC Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE.
RX MEDLINE=89008313; PubMed=3139658;
RA Mori M., Sakagami Y., Ishii Y., Isogai A., Kitada C., Fujino M.,
RA Adsit J.C., Dunny G.M., Suzuki A.;
RT "Structure of cCF10, a peptide sex pheromone which induces
RT conjugative transfer of the Streptococcus faecalis tetracycline
RT resistance plasmid, pCF10.";
RL J. Biol. Chem. 263:14574-14578(1988).
CC -!- FUNCTION: CCF10 IS INVOLVED IN THE CONJUGATIVE TRANSFER OF THE
CC HEMOLYSIN PLASMID PCF10.
DR PIR; A30812; A30812.
KW Pheromone.
SQ SEQUENCE 7 AA; 790 MW; 72C9D2C731B2C740 CRC64;

Query Match 22.2%; Score 10; DB 1; Length 7;
Best Local Similarity 50.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      4 TFYF 7
          |  |
Db      3 TLVF 6
```

Search completed: July 8, 2002, 11:42:02
Job time: 194 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 8, 2002, 11:38:28 ; Search time 23.86 Seconds
(without alignments)
58.003 Million cell updates/sec

Title: US-09-461-061A-1
Perfect score: 45
Sequence: 1 NNATFYFK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 378

Minimum DB seq length: 0
Maximum DB seq length: 8

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	19	42.2	8	8	Q9TD02	Q9td02 terranatos
2	16	35.6	8	10	Q40530	Q40530 nicotiana t
3	15	33.3	8	6	Q95M23	Q95m23 sus scrofa
4	15	33.3	8	11	Q9ERD2	Q9erd2 mus musculu
5	14	31.1	8	2	Q9R9C2	Q9r9c2 borrelia bu
6	14	31.1	8	8	Q94VF6	Q94vf6 varanus job
7	13	28.9	8	2	O09258	O09258 synechococc
8	13	28.9	8	2	O87471	O87471 haemophilus
9	13	28.9	8	5	Q9TWH6	Q9twh6 perinereis
10	12	26.7	8	2	O52062	O52062 bacillus me
11	12	26.7	8	2	Q49534	Q49534 mycoplasma
12	12	26.7	8	5	P82686	P82686 periplaneta
13	12	26.7	8	8	Q34909	Q34909 locusta mig
14	12	26.7	8	10	Q9SB24	Q9sb24 nicotiana t
15	12	26.7	8	10	Q40659	Q40659 oryza sativ
16	12	26.7	8	11	Q62933	Q62933 rattus norv
17	11	24.4	7	4	Q15903	Q15903 homo sapien
18	11	24.4	8	2	Q9R5L7	Q9r5l7 clostridium
19	11	24.4	8	2	Q9S6D5	Q9s6d5 escherichia
20	11	24.4	8	2	Q9S443	Q9s443 pseudomonas
21	11	24.4	8	2	Q45889	Q45889 clostridium
22	11	24.4	8	3	P87225	P87225 saccharomyc
23	11	24.4	8	4	Q9UJ50	Q9uj50 homo sapien
24	11	24.4	8	5	O15896	O15896 babesia bov
25	11	24.4	8	6	P82929	P82929 bos taurus
26	11	24.4	8	11	Q9R1U6	Q9r1u6 mus musculu
27	11	24.4	8	11	O35835	O35835 rattus sp.
28	10	22.2	7	2	O50556	O50556 actinobacil
29	10	22.2	7	2	Q47505	Q47505 escherichia
30	10	22.2	7	12	Q65578	Q65578 bovine herp
31	10	22.2	8	2	Q934S4	Q934s4 thiobacillu
32	10	22.2	8	3	O13591	O13591 saccharomyc
33	10	22.2	8	4	Q9HCQ0	Q9hcq0 homo sapien
34	10	22.2	8	4	Q9Y4J3	Q9y4j3 homo sapien
35	10	22.2	8	8	Q9TKE5	Q9tke5 leptospermu
36	10	22.2	8	8	Q9T4Y2	Q9t4y2 asterina pe
37	10	22.2	8	11	Q9QVJ8	Q9qvj8 mus sp. mep
38	10	22.2	8	11	Q99NX9	Q99nx9 hydrochoeru
39	10	22.2	8	12	Q91U21	Q91u21 influenza a
40	10	22.2	8	12	Q91U19	Q91u19 influenza a
41	9	20.0	5	13	P82070	P82070 litoria rub
42	9	20.0	5	13	P82071	P82071 litoria rub
43	9	20.0	5	13	P82072	P82072 litoria rub
44	9	20.0	5	13	P82073	P82073 litoria rub
45	9	20.0	8	3	P82858	P82858 puccinia re

ALIGNMENTS

RESULT 1
Q9TD02
ID Q9TD02 PRELIMINARY; PRT; 8 AA.
AC Q9TD02;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE CYTOCHROME C OXIDASE SUBUNIT I (FRAGMENT).
OS Terranatos dolichopterus.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Cyprinodontiformes; Aplocheilidae; Rivulinae; Terranatos.
OX NCBI_TaxID=61836;
RN [1]
RP SEQUENCE FROM N.A.
RA Hrbek T., Larson A.;
RT "The evolution of diapause in the killifish family Rivulidae
RT (Atherinomorpha, Cyprinodontiformes): A molecular phylogenetic and
RT biogeographic perspective.";
RL Evolution 53:1200-1216(1999).
DR EMBL; AF092421; AAF03041.1; -.
KW Mitochondrion.
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 1084 MW; F0C9D3640DD44056 CRC64;

Query Match 42.2%; Score 19; DB 8; Length 8;
Best Local Similarity 50.0%; Pred. No. 5.6e+05;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy      2 NATFYF 7
          || ::|
Db      3 NARWFF 8
```

RESULT 2
Q40530

```
ID   Q40530      PRELIMINARY;      PRT;     8 AA.
AC   Q40530;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   P20 N WITH A LEADER PEPTIDE.
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87089808; PubMed=3540612;
RA   Herman L.M.F., Montagu M.C.V., Depicker A.G.;
RT   "Isolation of tobacco DNA segments with plant promoter activity.";
RL   Mol. Cell. Biol. 6:4486-4492(1986).
DR   EMBL; M14685; AAA34090.1; -.
SQ   SEQUENCE   8 AA;  1109 MW;  E257205B19C9C9C6 CRC64;


  Query Match             35.6%;  Score 16;  DB 10;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 FYFK 8
          |:|:
Db      2 FFFE 5


RESULT   3
Q95M23
ID   Q95M23      PRELIMINARY;      PRT;     8 AA.
AC   Q95M23;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NA+/K+ ATPASE ALPHA SUBUNIT (EC 3.6.1.37) (FRAGMENT).
GN   ATP1A1.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PIETRAIN;
RA   Blazkova P., Stratil A., Peelman L.J., Van Poucke M., Reiner G.,
RA   Geldermann H., Kopecny M.;
RT   "RH mapping of the porcine ATP1A1, ATP1B1, V-ATPase, IVL genes and
RT   linkage assignments of ATP1A1 and IVL to chromosome 4.";
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ344138; CAC51422.1; -.
KW   Hydrolase.
FT   NON_TER       1      1
SQ   SEQUENCE   8 AA;  1117 MW;  604B41AB133B02D3 CRC64;


  Query Match             33.3%;  Score 15;  DB 6;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 TFY 6
          |:|
Db      6 TYY 8


RESULT   4
Q9ERD2
ID   Q9ERD2      PRELIMINARY;      PRT;     8 AA.
AC   Q9ERD2;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   GALACTOSE-1-PHOSPHATE URIDYL TRANSFERASE (FRAGMENT).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/D3;
RA   Leslie N.D., Bai S.;
RT   "Functional analysis of the mouse galactose-1-phosphate uridyl
RT   transferase (GALT) promoter.";
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF314226; AAG31161.1; -.
KW   Transferase.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  854 MW;  ECBDC409D1ADDDD6 CRC64;


  Query Match             33.3%;  Score 15;  DB 11;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 5.6e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 ATF 5
          |||
Db      3 ATF 5


RESULT   5
Q9R9C2
ID   Q9R9C2      PRELIMINARY;      PRT;     8 AA.
AC   Q9R9C2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PLASMID CP32-3, POSSIBLE PARTITION PROTEINS (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OG   Plasmid cp32-3.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B31;
RX   MEDLINE=98361033; PubMed=9695920;
RA   Stevenson B., Casjens S., Rosa P.;
RT   "Evidence of past recombination events among the genes encoding the
RT   Erp antigens of Borrelia burgdorferi.";
RL   Microbiology 144:1869-1879(1998).
DR   EMBL; AF022480; AAC35445.1; -.
KW   Plasmid.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  985 MW;  E8B41B41A735B446 CRC64;


  Query Match             31.1%;  Score 14;  DB 2;  Length 8;
  Best Local Similarity   40.0%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 NNATF 5
          |: |:
Db      2 NSLTY 6


RESULT   6
Q94VF6
ID   Q94VF6      PRELIMINARY;      PRT;     8 AA.
AC   Q94VF6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CYTOCHROME C OXIDASE SUBUNIT I (FRAGMENT).
GN   COI.
OS   Varanus jobiensis.
OG   Mitochondrion.
```

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
OX   NCBI_TaxID=169843;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ast J.C.;
RT   "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
RL   Cladistics 17:0-0(2001).
DR   EMBL; AF407507; AAL10075.1; -.
KW   Mitochondrion.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  1144 MW;  EFD729DB436411A6 CRC64;


  Query Match             31.1%;  Score 14;  DB 8;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 FYF 7
          :||
Db      4 WYF 6


RESULT   7
O09258
ID   O09258      PRELIMINARY;      PRT;     8 AA.
AC   O09258;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NIFH (FRAGMENT).
GN   NIFH.
OS   Synechococcus sp. (strain PCC 8801 / RF-1) (Cyanothece PCC 8801).
OC   Bacteria; Cyanobacteria; Chroococcales; Cyanothece.
OX   NCBI_TaxID=41431;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RF-1;
RX   MEDLINE=99231861; PubMed=10217509;
RA   Huang T.C., Lin R.F., Chu M.K., Chen H.M.;
RT   "Organization and expression of nitrogen-fixation genes in the aerobic
RT   nitrogen-fixing unicellular cyanobacterium Synechococcus sp. strain
RT   RF-1.";
RL   Microbiology 145:743-753(1999).
DR   EMBL; AF001780; AAC33369.1; -.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  985 MW;  F16B59CDD046C406 CRC64;


  Query Match             28.9%;  Score 13;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 FY 6
          ||
Db      6 FY 7


RESULT   8
O87471
ID   O87471      PRELIMINARY;      PRT;     8 AA.
AC   O87471;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   HIFA (FRAGMENT).
GN   HIFA.
OS   Haemophilus influenzae.
OC   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC   Haemophilus.
OX   NCBI_TaxID=727;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EAGAN (E1A);
RX   MEDLINE=98389689; PubMed=9721313;
RA   Mhlanga-Mutangadura T., Morlin G., Smith A.L., Eisenstark A.,
RA   Golomb M.;
RT   "Evolution of the major pilus gene cluster of haemophilus
RT   influenzae.";
RL   J. Bacteriol. 180:4693-4703(1998).
DR   EMBL; AF071762; AAC35830.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   8 AA;  876 MW;  DAB44451A7272325 CRC64;


  Query Match             28.9%;  Score 13;  DB 2;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 NNA 3
          ||:
Db      6 NNS 8


RESULT   9
Q9TWH6
ID   Q9TWH6      PRELIMINARY;      PRT;     8 AA.
AC   Q9TWH6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   BIOACTIVE PEPTIDE P4=PUTATIVE ESOPHAGEAL NEUROREGULATOR.
OS   Perinereis vancaurica.
OC   Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Aciculata;
OC   Phyllodocida; Nereididae; Perinereis.
OX   NCBI_TaxID=6355;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=95323338; PubMed=7599979;
RA   Takahashi T., Furukawa Y., Muneoka Y., Matsushima O., Ikeda T.,
RA   Fujita T., Minakata H., Nomoto K.;
RT   "Isolation and characterization of four novel bioactive peptides from
RT   a polychaete annelid, Perinereis vancaurica.";
RL   Comp. Biochem. Physiol. C,
RL   Pharmacol. Toxicol. Endocrinol. 110:297-304(1995).
SQ   SEQUENCE   8 AA;  989 MW;  954772CAA87B0B59 CRC64;


  Query Match             28.9%;  Score 13;  DB 5;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 FY 6
          ||
Db      1 FY 2


RESULT  10
O52062
ID   O52062      PRELIMINARY;      PRT;     8 AA.
AC   O52062;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   SIGMA FACTOR (FRAGMENT).
GN   SPOIIGB.
OS   Bacillus megaterium.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1404;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Fawcett P., Melnikov A., Youngman P.;
RT   "The Bacillus SpoIIGA protein is targeted to sites of sporulation
RT   septum formation in a SpoIIE independent manner.";
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
```

```
DR   EMBL; AF017181; AAB94056.1; -.
DR   MEROPS; U04.001; -.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  1057 MW;  F0D400441B41ADD6 CRC64;


  Query Match             26.7%;  Score 12;  DB 2;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 ATF 5
          ||:
Db      2 ATY 4


RESULT  11
Q49534
ID   Q49534      PRELIMINARY;      PRT;     8 AA.
AC   Q49534;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   P120 (FRAGMENT).
GN   P120.
OS   Mycoplasma hominis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC   Mycoplasmataceae; Mycoplasma.
OX   NCBI_TaxID=2098;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=V2785;
RA   Nyvold C., Birkelund S., Christiansen G.;
RT   "The Mycoplasma hominis P120 membrane protein gene contains a 659 base
RT   pair hypervariable domain.";
RL   Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U22025; AAA67455.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   8 AA;  869 MW;  914457605B02C05D CRC64;


  Query Match             26.7%;  Score 12;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 NN 2
          ||
Db      7 NN 8


RESULT  12
P82686
ID   P82686      PRELIMINARY;      PRT;     8 AA.
AC   P82686;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   KININ-2 (PEA-K-2).
OS   Periplaneta americana (American cockroach).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC   Blattoidea; Blattidae; Periplaneta.
OX   NCBI_TaxID=6978;
RN   [1]
RP   SEQUENCE, AND FUNCTION.
RC   TISSUE=CORPORA CARDIACA;
RX   MEDLINE=98010462; PubMed=9350979;
RA   Predel R., Kellner R., Rapus J., Penzlin H., Gaede G.;
RT   "Isolation and structural elucidation of eight kinins from the
RT   retrocerebral complex of the American cockroach, Periplaneta
RT   americana.";
RL   Regul. Pept. 71:199-205(1997).
CC   -!- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
CC       (MYOTROPIC ACTIVITY).
CC   -!- SIMILARITY: BELONGS TO THE KININ FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  856 MW;  DC6365A5B9D5BDDA CRC64;


  Query Match             26.7%;  Score 12;  DB 5;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 NATF 5
          :|:|
Db      1 DASF 4


RESULT  13
Q34909
ID   Q34909      PRELIMINARY;      PRT;     8 AA.
AC   Q34909;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CYTOCHROME B (FRAGMENT).
OS   Locusta migratoria (Migratory locust).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC   Acridomorpha; Acridoidea; Acrididae; Locusta.
OX   NCBI_TaxID=7004;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88223478; PubMed=2836084;
RA   McCracken A., Uhlenbusch I., Gellissen G.;
RT   "Structure of the cloned Locusta migratoria mitochondrial genome:
RT   restriction mapping and sequence of its ND-1(URF-1) gene.";
RL   Curr. Genet. 11:625-630(1987).
DR   EMBL; X05286; CAA28905.1; -.
KW   Mitochondrion.
FT   NON_TER       1      1
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  1019 MW;  F8E33723304B45B6 CRC64;


  Query Match             26.7%;  Score 12;  DB 8;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      6 YFK 8
          | |
Db      3 YIK 5


RESULT  14
Q9SB24
ID   Q9SB24      PRELIMINARY;      PRT;     8 AA.
AC   Q9SB24;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   PARAT PROTEIN (FRAGMENT).
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Niwa Y., Muranaka T., Baba A., Machida Y.;
RT   "Organ-specific and auxin-inducible expression of two tobacco parA-
RT   related genes in transgenic plants.";
RL   DNA Res. 0:0-0(1994).
DR   EMBL; D42119; BAA07700.1; -.
FT   NON_TER       8      8
```

```
SQ   SEQUENCE   8 AA;  905 MW;  FE32D2C44455BB16 CRC64;

  Query Match             26.7%;  Score 12;  DB 10;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 NN 2
          ||
Db      4 NN 5


RESULT  15
Q40659
ID   Q40659      PRELIMINARY;      PRT;     8 AA.
AC   Q40659;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ALPHA-AMYLASE (FRAGMENT).
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91078641; PubMed=2258052;
RA   Kumagai M.H., Shah M., Terashima M., Vrkljan Z., Whitaker J.R.,
RA   Rodriquez R.L.;
RT   "Expression and secretion of rice alpha-amylase by saccharomycers
RT   cerevisiae.";
RL   Gene 94:209-216(1990).
DR   EMBL; M62916; AAA33892.1; -.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  948 MW;  EBC694444732D6D6 CRC64;

  Query Match             26.7%;  Score 12;  DB 10;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 NN 2
          ||
Db      5 NN 6


RESULT  16
Q62933
ID   Q62933      PRELIMINARY;      PRT;     8 AA.
AC   Q62933;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HISTONE H10 (FRAGMENT).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR; TISSUE=TESTIS;
RA   Khochbin S., Peretti M.F., Girardot V.;
RT   "The structure of differentiation-specific histone H10 gene basal
RT   promoter.";
RL   Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U49737; AAA92724.1; -.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  850 MW;  F4C771B5B44B01A6 CRC64;

  Query Match             26.7%;  Score 12;  DB 11;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 NAT 4
          |:|
Db      4 NST 6


RESULT  17
Q15903
ID   Q15903      PRELIMINARY;      PRT;     7 AA.
AC   Q15903;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   (CLONE XP7E7B) (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA;
RA   Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA   Coolbaugh M.I., Chinault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,
RA   Caskey C.T.H.;
RT   "Isolation of chromosome-specific genes by reciprocal probing of
RT   arrayed cDNAs and cosmid libraries.";
RL   Hum. Mol. Genet. 0:0-0(1995).
DR   EMBL; L32082; AAA73893.1; -.
FT   NON_TER       1      1
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  849 MW;  6B040339CDD33DB0 CRC64;

  Query Match             24.4%;  Score 11;  DB 4;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      7 FK 8
          ||
Db      4 FK 5


RESULT  18
Q9R5L7
ID   Q9R5L7      PRELIMINARY;      PRT;     8 AA.
AC   Q9R5L7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   1,4-BETA-D-GLUCAN GLUCANOHYDROLASE (EC 3.2.1.4) (FRAGMENT).
OS   Clostridium thermocellum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1515;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92231850; PubMed=1567379;
RA   Romaniec M.P., Fauth U., Kobayashi T., Huskisson N.S., Barker P.J.,
RA   Demain A.L.;
RT   "Purification and characterization of a new endoglucanase from
RT   Clostridium thermocellum.";
RL   Biochem. J. 283:69-73(1992).
SQ   SEQUENCE   8 AA;  823 MW;  C2C1AB1DD9D1B775 CRC64;

  Query Match             24.4%;  Score 11;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 TF 5
          ||
Db      3 TF 4
```

```
RESULT 19
Q9S6D5
ID   Q9S6D5      PRELIMINARY;      PRT;     8 AA.
AC   Q9S6D5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   PUTATIVE IS30 TRANSPOSASE (FRAGMENT).
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A295B;
RX   MEDLINE=99194747; PubMed=10094716;
RA   Rahn A., Drummelsmith J., Whitfield C.;
RT   "Conserved organization in the cps gene clusters for expression of
RT   Escherichia coli group 1 K antigens: relationship to the colanic acid
RT   biosynthesis locus and the cps genes from Klebsiella pneumoniae.";
RL   J. Bacteriol. 181:2307-2313(1999).
DR   EMBL; AF118251; AAD30008.1; -.
FT   NON_TER       8       8
SQ   SEQUENCE   8 AA;  1011 MW;  F21DC1A9D1B41406 CRC64;


  Query Match             24.4%;  Score 11;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 TF 5
          ||
Db      4 TF 5


RESULT 20
Q9S443
ID   Q9S443      PRELIMINARY;      PRT;     8 AA.
AC   Q9S443;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   BETA-LACTAMASE (FRAGMENT).
GN   PSE2.
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R545;
RA   Roy D., Coulombe M., Perron K., Roy P.H.;
RT   "Characterization of a novel 6'-N-aminoglycoside acetyltransferase
RT   gene aac(6')-IIc from the integron of a Chinese Pseudomonas aeruginosa
RT   clinical isolate.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF162771; AAD46628.1; -.
FT   NON_TER       8       8
SQ   SEQUENCE   8 AA;  930 MW;  EBDB5DDDD9D1A336 CRC64;


  Query Match             24.4%;  Score 11;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 TF 5
          ||
Db      3 TF 4
```

Search completed: July 8, 2002, 11:41:46
Job time: 198 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 8, 2002, 11:37:53 ; Search time 14.17 Seconds
(without alignments)
54.249 Million cell updates/sec

Title: US-09-461-061A-1
Perfect score: 45
Sequence: 1 NNATFYFK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 603

Minimum DB seq length: 0
Maximum DB seq length: 8

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	16	35.6	8	2	C61512	variant surface gl
2	16	35.6	8	2	JS0318	leucokinin VIII -
3	15	33.3	7	2	A34026	acetylcholinestera
4	15	33.3	7	2	E33932	Ig mu chain D regi
5	14	31.1	4	2	JQ1273	neuropeptide Antho
6	14	31.1	7	2	S17976	glucose isomerase
7	14	31.1	8	2	B27867	homeotic protein U
8	13	28.9	3	3	S68328	blood cell protein
9	13	28.9	5	2	PT0278	Ig heavy chain CRD
10	13	28.9	8	2	B45800	serum albumin - do
11	13	28.9	8	2	S68325	blood cell protein
12	13	28.9	8	2	A59028	MHC class I histoc
13	12	26.7	5	2	S51077	alpha-amylase - ri
14	12	26.7	5	4	A58728	serrawettin W2 - S
15	12	26.7	8	2	B24749	neuropeptide B - b
16	12	26.7	8	2	S70727	ipgF protein - Shi
17	12	26.7	8	2	A14683	aspartate transami
18	11	24.4	4	2	S53508	starvation-induced
19	11	24.4	5	2	S70615	endo-1,4-beta-xyla
20	11	24.4	6	2	S11024	hydrogensulfite re
21	11	24.4	6	2	PT0650	T-cell receptor be
22	11	24.4	6	2	A43129	neuropeptide GNFFR
23	11	24.4	7	2	S21230	dermorphin (Trp-4,
24	11	24.4	7	2	B35890	RNA-directed DNA p
25	11	24.4	7	2	S33567	tubulin beta-3 cha
26	11	24.4	7	2	PD0029	pev-kinin 1 - pena
27	11	24.4	8	2	S08996	hypertrehalosemic
28	11	24.4	8	2	B49823	adipokinetic hormo
29	11	24.4	8	2	B44960	neuropeptide Led-C
30	11	24.4	8	2	A33995	adipokinetic hormo
31	11	24.4	8	2	S11545	adipokinetic hormo
32	11	24.4	8	2	S37141	rpsA protein - Erw
33	11	24.4	8	2	S21273	cellulase (EC 3.2.
34	11	24.4	8	2	S13661	polygalacturonase
35	11	24.4	8	2	A39892	P element, P cytot
36	11	24.4	8	2	PH0803	T-cell receptor al
37	11	24.4	8	2	S65381	cytochrome-c oxida
38	11	24.4	8	2	PC4373	telomeric and tetr
39	10	22.2	4	2	S43014	hypothetical prote
40	10	22.2	5	2	F22565	R-phycoerythrin ga
41	10	22.2	5	2	JH0253	gut pentapeptide -
42	10	22.2	6	2	A61411	ameletin - rat
43	10	22.2	6	2	I51317	bHLH transcription
44	10	22.2	6	2	PT0518	T-cell receptor be
45	10	22.2	6	2	PT0587	T-cell receptor be

ALIGNMENTS

RESULT 1
C61512
variant surface glycoprotein MITat 1.4 - Trypanosoma brucei (fragment)
C;Species: Trypanosoma brucei
C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 07-May-1999
C;Accession: C61512
R;Holder, A.A.; Cross, G.A.M.
Mol. Biochem. Parasitol. 2, 135-150, 1981
A;Title: Glycopeptides from variant surface glycoproteins of Trypanosoma brucei. C-te
A;Reference number: A61512; MUID:81172836
A;Accession: C61512
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-8 <HOL>
C;Keywords: glycoprotein

Query Match 35.6%; Score 16; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NNA 3
|||
Db 3 NNA 5

RESULT 2
JS0318
leucokinin VIII - Madeira cockroach
C;Species: Leucophaea maderae (Madeira cockroach)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Jun-2000
C;Accession: JS0318
R;Holman, G.M.; Cook, B.J.; Nachman, R.J.
Comp. Biochem. Physiol. C 88, 31-34, 1987
A;Title: Isolation, primary structure and synthesis of leucokinins VII and VIII: the
A;Reference number: JS0317
A;Accession: JS0318
A;Molecule type: protein
A;Residues: 1-8 <HOL>
C;Comment: Leucokinins, a family of cephalomyotropic peptides, stimulate contractile
C;Keywords: amidated carboxyl end; cephalomyotropic peptide
F;8/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 35.6%; Score 16; DB 2; Length 8;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 ATFY 6
| ||
Db 2 ADFY 5

RESULT 3
A34026
acetylcholinesterase (EC 3.1.1.7) 5.6S form - Pacific electric ray (fragment)
C;Species: Torpedo californica (Pacific electric ray)
C;Date: 16-Mar-1990 #sequence_revision 16-Mar-1990 #text_change 08-Nov-1996
C;Accession: A34026
R;Gibney, G.; MacPhee-Quigley, K.; Thompson, B.; Vedvick, T.; Low, M.G.; Taylor, S.S.; T
J. Biol. Chem. 263, 1140-1145, 1988
A;Title: Divergence in primary structure between the molecular forms of acetylcholineste
A;Reference number: A34026; MUID:88087239
A;Accession: A34026
A;Molecule type: protein
A;Residues: 1-7 <GIB>
C;Keywords: alternative splicing; carboxylic ester hydrolase

Query Match 33.3%; Score 15; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NAT 4
|||
Db 3 NAT 5

RESULT 4
E33932
Ig mu chain D region (E7) - mouse
C;Species: Mus musculus (house mouse)
C;Date: 09-Mar-1990 #sequence_revision 09-Mar-1990 #text_change 16-Aug-1996
C;Accession: E33932
R;Baccala, R.; Vo Quang, T.; Gilbert, M.; Ternynck, T.; Avrameas, S.
Proc. Natl. Acad. Sci. U.S.A. 86, 4624-4628, 1989
A;Title: Two murine natural polyreactive autoantibodies are encoded by nonmutated germ-l
A;Reference number: A33932; MUID:89282823
A;Accession: E33932
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-7 <BAC>
A;Cross-references: GB:M27106
C;Keywords: immunoglobulin

Query Match 33.3%; Score 15; DB 2; Length 7;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 FYFK 8
|| |
Db 3 FYGK 6

RESULT 5
JQ1273
neuropeptide Antho-KAamide - sea anemone (Anthopleura elegantissima)
C;Species: Anthopleura elegantissima
C;Date: 31-Mar-1992 #sequence_revision 04-Dec-1992 #text_change 08-Dec-1995
C;Accession: JQ1273
R;Nothacker, H.P.; Rinehart, K.L.; Grimmelikhuijzen, C.J.P.
Biochem. Biophys. Res. Commun. 179, 1205-1211, 1991
A;Title: Isolation of L-3-phenyllactyl-Phe-Lys-Ala-NH2 (Antho-KAamide), a novel neuropep
A;Reference number: JQ1273; MUID:92028852
A;Accession: JQ1273
A;Molecule type: protein
A;Residues: 1-4 <NOT>
C;Comment: The carboxyl-terminal amide probably arises from cleavage of a following glyc
C;Keywords: amidated carboxyl end; neuropeptide; phenyllactylation
F;1/Modified site: L-3-phenyllactic acid (Phe) #status experimental
F;4/Modified site: amidated carboxyl end (Ala) #status experimental

Query Match 31.1%; Score 14; DB 2; Length 4;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 6 YFK 8
:||
Db 1 FFK 3

RESULT 6
S17976
glucose isomerase - Thermoanaerobacterium saccharolyticum (fragment)
C;Species: Thermoanaerobacterium saccharolyticum
C;Date: 12-Feb-1998 #sequence_revision 12-Feb-1998 #text_change 17-Apr-1998
C;Accession: S17976
R;Lee, C.; Zeikus, J.G.
Biochem. J. 273, 565-571, 1991
A;Title: Purification and characterization of thermostable glucose isomerase from Clo
A;Reference number: S15119; MUID:91144536
A;Accession: S17976
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <LEE>
A;Note: the sequence from page 568 is inconsistent with that from page 565 in having
A;Note: the source is designated as Thermoanaerobacter strain B6A

Query Match 31.1%; Score 14; DB 2; Length 7;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 6 YFK 8
||:
Db 4 YFE 6

RESULT 7
B27867
homeotic protein Ultrabithorax - fruit fly (Drosophila melanogaster) (fragment)
C;Species: Drosophila melanogaster
C;Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 23-Feb-1997
C;Accession: B27867
R;Saari, G.; Bienz, M.
EMBO J. 6, 1775-1779, 1987
A;Title: The structure of the ultrabithorax promoter of Drosophila melanogaster.
A;Reference number: A91072
A;Accession: B27867
A;Molecule type: mRNA
A;Residues: 1-8 <SAA>
C;Genetics:
A;Gene: FlyBase:Ubx
A;Cross-references: FlyBase:FBgn0003944
C;Keywords: DNA binding; nucleus; transcription regulation

Query Match 31.1%; Score 14; DB 2; Length 8;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 6 YFK 8
||:
Db 4 YFE 6

RESULT 8
S68328
blood cell protein A - Molgula manhattensis (fragment)
C;Species: Molgula manhattensis
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C;Accession: S68328
R;Taylor, S.W.; Ross, M.M.; Waite, J.H.
Arch. Biochem. Biophys. 324, 228-240, 1995
A;Title: Novel 3,4-di- and 3,4,5-trihydroxyphenylalanine-containing polypeptides from
A;Reference number: S68325; MUID:96132650

A;Accession: S68328
A;Molecule type: protein
A;Residues: 1-3 <TAY>

Query Match 28.9%; Score 13; DB 3; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FY 6
||
Db 2 FY 3

RESULT 9
PT0278
Ig heavy chain CRD3 region (clone 4-88) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PT0278
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and j
A;Reference number: PT0222; MUID:91108337
A;Accession: PT0278
A;Molecule type: DNA
A;Residues: 1-5 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotetramer; immunoglobulin

Query Match 28.9%; Score 13; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 YF 7
||
Db 1 YF 2

RESULT 10
B45800
serum albumin - dog (fragment)
C;Species: Canis lupus familiaris (dog)
C;Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 31-Dec-1993
C;Accession: B45800
R;Carraway, R.E.; Cochrane, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A;Title: Structures of histamine-releasing peptides formed by the action of acid proteas
A;Reference number: A45800; MUID:89341406
A;Accession: B45800
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-8 <CAR>

Query Match 28.9%; Score 13; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 YF 7
||
Db 7 YF 8

RESULT 11
S68325
blood cell protein A8 - Ascidia ceratodes (fragment)
N;Alternate names: Abcp-A
C;Species: Ascidia ceratodes
C;Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 11-Jun-1999
C;Accession: S68325
R;Taylor, S.W.; Ross, M.M.; Waite, J.H.
Arch. Biochem. Biophys. 324, 228-240, 1995
A;Title: Novel 3,4-di- and 3,4,5-trihydroxyphenylalanine-containing polypeptides from
A;Reference number: S68325; MUID:96132650
A;Accession: S68325
A;Molecule type: protein
A;Residues: 1-8 <TAY>
F;2,3,4/Modified site: 3',4',5'-trihydroxyphenylalanine (Tyr) #status experimental

Query Match 28.9%; Score 13; DB 2; Length 8;
Best Local Similarity 33.3%; Pred. No. 2.8e+05;
Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FYF 7
:|:
Db 2 YYY 4

RESULT 12
A59028
MHC class I histocompatibility antigen HLA-C alpha chain - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 31-Dec-2001 #sequence_revision 31-Dec-2001 #text_change 31-Dec-2001
C;Accession: A59028
R;Escolano, J.M.
submitted to the Protein Sequence Database, August 1998
A;Reference number: A59028
A;Accession: A59028
A;Molecule type: protein
A;Residues: 1-8 <ESC>
C;Keywords: glycoprotein; heterodimer; transmembrane protein; transplantation antigen

Query Match 28.9%; Score 13; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 YF 7
||
Db 1 YF 2

RESULT 13
S51077
alpha-amylase - rice
C;Species: Oryza sativa (rice)
C;Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 01-Sep-1995
C;Accession: S51077
R;Terashima, M.; Kubo, A.; Suzawa, M.; Itoh, Y.; Katoh, S.
Eur. J. Biochem. 226, 249-254, 1994
A;Title: The roles of the N-linked carbohydrate chain of rice alpha-amylase in thermo
A;Reference number: S51077; MUID:95045597
A;Accession: S51077
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-5 <TER>

Query Match 26.7%; Score 12; DB 2; Length 5;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NNAT 4
:| |
Db 1 SNGT 4

RESULT 14
A58728
serrawettin W2 - Serratia marcescens
C;Species: Serratia marcescens
C;Date: 10-Feb-1998 #sequence_revision 12-Feb-1998 #text_change 12-Feb-1998
C;Accession: A58728

```
R;Matsuyama, T.; Kaneda, K.; Nakagawa, Y.; Isa, K.; Hara-Hotta, H.; Yano, I.
J. Bacteriol. 174, 1769-1776, 1992
A;Title: A novel extracellular cyclic lipopeptide which promotes flagellum-dependent and
A;Reference number: A58728; MUID:92193260
A;Accession: A58728
A;Status: unencoded polypeptide
A;Molecule type: protein
A;Residues: 1-5 <MAT>
A;Experimental source: strain NS 25
C;Comment: A surfactant lipopeptide promoting flagellum-independent surface translocatic
C;Keywords: blocked amino end; blocked carboxyl end; D-amino acid; lipoprotein; unencode
F;1/Modified site: D-leucine (Leu) #status experimental
F;4/Modified site: D-phenylalanine (Phe) #status experimental
F;1-5/Cross-link: 3-hydroxydecanoyl amino end, ester carboxyl end (Leu-Ile) #status expe


  Query Match             26.7%;  Score 12;  DB 4;  Length 5;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 ATF 5
           :||
Db       2 STF 4


RESULT  15
B24749
neuropeptide B - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 28-Jul-1987 #sequence_revision 28-Jul-1987 #text_change 18-Aug-2000
C;Accession: B24749
R;Yang, H.Y.T.; Fratta, W.; Majane, E.A.; Costa, E.
Proc. Natl. Acad. Sci. U.S.A. 82, 7757-7761, 1985
A;Title: Isolation, sequencing, synthesis, and pharmacological characterization of two b
A;Reference number: A94074; MUID:86067985
A;Accession: B24749
A;Molecule type: protein
A;Residues: 1-8 <YAN>
C;Superfamily: unassigned animal peptides
C;Keywords: neuropeptide


  Query Match             26.7%;  Score 12;  DB 2;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 FYFK 8
           | |:
Db       1 FLFQ 4


RESULT  16
S70727
ipgF protein - Shigella flexneri (fragment)
C;Species: Shigella flexneri
C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 08-Oct-1999
C;Accession: S70727
R;Allaoui, A.; Sansonetti, P.J.; Menard, R.; Barzu, S.; Mounier, J.; Phalipon, A.; Parsc
Mol. Microbiol. 17, 461-470, 1995
A;Title: MxiG, a membrane protein required for secretion of Shigella spp. Ipa invasins:
A;Reference number: S70727; MUID:96100445
A;Accession: S70727
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-8 <ALL>
A;Cross-references: EMBL:Z48957; NID:g929880; PIDN:CAA88821.1; PID:g929881
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1995
C;Genetics:
A;Gene: ipgF


  Query Match             26.7%;  Score 12;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NN 2
             ||
Db       4 NN 5


RESULT  17
A14683
aspartate transaminase (EC 2.6.1.1), mitochondrial - chicken (tentative sequence) (fr
N;Alternate names: aspartate aminotransferase, mitochondrial
C;Species: Gallus gallus (chicken)
C;Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 31-Mar-2000
C;Accession: A14683
R;Wilson, K.J.; Hunziker, P.; Hughes, G.J.
FEBS Lett. 108, 98-102, 1979
A;Title: Microsequence analysis: IV. Automatic liquid-phase sequencing using DABITC.
A;Reference number: A14683; MUID:80092116
A;Accession: A14683
A;Molecule type: protein
A;Residues: 1-8 <WIL>
C;Keywords: aminotransferase; mitochondrion


  Query Match             26.7%;  Score 12;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NN 2
             ||
Db       3 NN 4


RESULT  18
S53508
starvation-induced ribonuclease - tomato
C;Species: Lycopersicon esculentum (tomato)
C;Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 07-May-1999
C;Accession: S53508
R;Koeck, M.; Loeffler, A.; Abel, S.; Glund, K.
Plant Mol. Biol. 27, 477-485, 1995
A;Title: cDNA structure and regulatory properties of a family of starvation-induced r
A;Reference number: S53506; MUID:95201242
A;Accession: S53508
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-4 <KOE>


  Query Match             24.4%;  Score 11;  DB 2;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 TF 5
             ||
Db       3 TF 4


RESULT  19
S70615
endo-1,4-beta-xylanase (EC 3.2.1.8) - Streptomyces sp. (Chainia sp. NCL 82.5.1) (frag
N;Alternate names: xylanase
C;Species: Streptomyces sp.
A;Variety: Chainia sp. NCL 82.5.1
C;Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 07-May-1999
C;Accession: S70615
R;Rao, M.; Khadilkar, S.; Bandivadekar, K.R.; Deshpande, V.
Biochem. J. 316, 771-775, 1996
A;Title: Structural environment of an essential cysteine residue of xylanase from Cha
A;Reference number: S70615; MUID:96265041
A;Accession: S70615
A;Molecule type: protein
A;Residues: 1-5 <RAO>
```

A;Experimental source: Chainia sp. strain NCL 82.5.1
A;Note: the source is designated as Chainia sp.
C;Function:
A;Description: endohydrolyzation of beta-1,4-xylosidic linkages in xylans
A;Pathway: fermentation of hemicellulose into ethanol
C;Keywords: glycosidase; hydrolase

Query Match 24.4%; Score 11; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TF 5
||
Db 2 TF 3

RESULT 20
S11024
hydrogensulfite reductase (EC 1.8.99.3) chain 1 - Desulfovibrio thermophilus (fragment)
N;Alternate names: bisulfite reductase; desulfofuscidin
C;Species: Desulfovibrio thermophilus
C;Date: 19-Mar-1997 #sequence_revision 30-Jan-1998 #text_change 30-Jan-1998
C;Accession: S11024
R;Fauque, G.; Lino, A.R.; Czechowski, M.; Kang, L.; DerVartanian, D.V.; Moura, J.J.G.; L
Biochim. Biophys. Acta 1040, 112-118, 1990
A;Title: Purification and characterization of bisulfite reductase (desulfofuscidin) from
A;Reference number: S11024; MUID:90335276
A;Accession: S11024
A;Molecule type: protein
A;Residues: 1-6 <FAU>
C;Keywords: oxidoreductase

Query Match 24.4%; Score 11; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 FK 8
||
Db 5 FK 6

Search completed: July 8, 2002, 11:39:07
Job time: 74 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 8, 2002, 11:37:48 ; Search time 29.04 Seconds
(without alignments)
30.599 Million cell updates/sec

Title: US-09-461-061A-1
Perfect score: 45
Sequence: 1 NNATFYFK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 69368

Minimum DB seq length: 0
Maximum DB seq length: 8

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	64.4	8	22	ABB55873	Vascular dementia-
2	29	64.4	8	22	AAU28595	DPI tryptic digest
3	29	64.4	8	22	AAU24794	Schizophrenia-Asso
4	29	64.4	8	22	AAU26242	Depression-Associa
5	29	64.4	8	22	AAU15138	Schizophrenia-asso
6	24	53.3	8	16	AAR73374	Human TSH receptor
7	23	51.1	6	22	AAU08641	Interleukin-1 inhi
8	23	51.1	7	20	AAY40730	S4 derivative #4,
9	23	51.1	7	21	AAB30068	Scaffold protein S
10	22	48.9	6	20	AAY23893	Human vascular end
11	22	48.9	6	20	AAW97492	Antigenic site of
12	22	48.9	7	20	AAY23895	Human vascular end
13	21	46.7	4	18	AAW20072	HOX5, a hexose oxi
14	21	46.7	5	17	AAW00431	Interleukin-6 anta
15	21	46.7	6	22	AAB04166	Internal peptide f
16	21	46.7	8	18	AAW11529	Interleukin-6 anta
17	20	44.4	5	19	AAW86298	Rodent IL-1 delta
18	20	44.4	6	20	AAY25842	Human secreted pro
19	20	44.4	6	22	AAU08625	Interleukin-1 inhi
20	20	44.4	7	17	AAW94661	GL-7ACA acylase al
21	20	44.4	7	20	AAY40732	S4 derivative #6,
22	20	44.4	7	20	AAY40733	S4 derivative #7,
23	20	44.4	7	21	AAB30070	Scaffold protein S
24	20	44.4	7	21	AAB30071	Scaffold protein S
25	20	44.4	7	22	AAU08638	Interleukin-1 inhi
26	20	44.4	8	14	AAR43431	Ro/SSA epitope 47.
27	20	44.4	8	15	AAR61722	PLP peptide 74, po
28	20	44.4	8	16	AAR71815	A.alternata allerg
29	20	44.4	8	22	AAE12635	Human gene 2 encod
30	20	44.4	8	22	AAE08103	Peptide #2 from hu
31	19	42.2	6	16	AAY03970	Cyclic somatostati
32	19	42.2	6	17	AAW08639	Met-enkephalin ana
33	19	42.2	6	17	AAW08636	Met-enkephalin ana
34	19	42.2	6	20	AAY55276	ATCC HB 11885 mono
35	19	42.2	6	21	AAB17213	IL-1 antagonist pe
36	19	42.2	6	21	AAY87022	Human haematopoiet
37	19	42.2	7	15	AAR44688	Somatostatin recep
38	19	42.2	7	16	AAY03971	Cyclic somatostati
39	19	42.2	7	16	AAY03972	Cyclic somatostati
40	19	42.2	7	16	AAR77965	Antigenic Tbp2 pep
41	19	42.2	7	17	AAR90409	Hybridoma ATCC HB-
42	19	42.2	7	18	AAW46158	Predicted antigeni
43	19	42.2	7	19	AAY17855	Backbone cyclised
44	19	42.2	7	19	AAW62500	Antibiotic peptide
45	19	42.2	7	19	AAW58754	SEQ ID NO:82 from

ALIGNMENTS

RESULT 1
ABB55873
ID ABB55873 standard; Peptide; 8 AA.
XX
AC ABB55873;
XX
DT 15-FEB-2002 (first entry)
XX
DE Vascular dementia-associated protein isoform (VPI) 73.
XX
KW Vascular Dementia; VD; VD-associated protein isoform; VPI; screening;
KW diagnosis; prognosis; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200169261-A2.
XX
PD 20-SEP-2001.
XX
PF 14-MAR-2001; 2001WO-GB01106.
XX
PR 15-MAR-2000; 2000GB-0006285.
PR 24-NOV-2000; 2000GB-0028734.
PR 28-NOV-2000; 2000US-0724391.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Herath HMAC, Parekh RB, Rohlff C;
XX
DR WPI; 2001-557937/62.
XX
PT Screening, diagnosis or prognosis of vascular dementia (VD), useful for
PT determining stage of VD and monitoring the effect of VD therapy,
PT comprises analysing body fluid by 2-dimensional electrophoresis for

PT features correlated with VD -
XX
PS Claim 6; Page 31; 151pp; English.
XX
CC The invention relates to screening, diagnosis or prognosis of Vascular
CC Dementia (VD) in a subject comprising analysing body fluid from the
CC subject by 2-dimensional (2-D) electrophoresis to generate a 2-D array of
CC features containing at least one chosen feature whose relative abundance
CC correlates with the presence, absence, stage or severity of VD or
CC predicts the onset or course of VD, especially detecting in a sample of
CC cerebrospinal fluid (CSF) from the subject one of 223 VD-associated
CC protein isoforms (VPIs) (ABB55801-ABB56295) as fully defined in the
CC specification. Detecting VD-associated features and VPI is useful for the
CC screening, diagnosis or prognosis of VD, for determining the stage or
CC severity of VD, for identifying a subject at risk of VD or for
CC monitoring the effect of therapy administered to a subject having VD.
CC Nucleic acids encoding a VPI or inhibiting the function of a VPI are
CC useful for the treatment of VD and for gene therapy.
XX
SQ Sequence 8 AA;

Query Match 64.4%; Score 29; DB 22; Length 8;
Best Local Similarity 75.0%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NNATFYFK 8
||| |||
Db 1 nnahgyfk 8

RESULT 2
AAU28595
ID AAU28595 standard; Peptide; 8 AA.
XX
AC AAU28595;
XX
DT 03-JAN-2002 (first entry)
XX
DE DPI tryptic digest peptide #192.
XX
KW Human; depression associated protein isoform; tryptic digest peptide;
KW DPI; cerebrospinal fluid; CSF; BAD; bipolar affective disorder;
KW neuropsychiatric disorder; bipolar mood disorder; neuroleptic;
KW maniac-depressive illness; schizoaffective disorder.
XX
OS Homo sapiens.
XX
PN WO200162787-A1.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-GB00786.
XX
PR 24-FEB-2000; 2000GB-0004412.
PR 08-DEC-2000; 2000GB-0030050.
PR 12-DEC-2000; 2000US-0254830.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Herath HMAC, Parekh RB, Rohlff C, Terrett JA, Tyson KL;
XX
DR WPI; 2001-570626/64.
XX
PT Novel nucleic acid encoding a protein associated with bipolar affective
PT disorder, which is used for diagnosis, prophylaxis and therapy of
PT neuropsychiatric disorders, such as bipolar affective disorder -
XX
PS Disclosure; Page 34; 153pp; English.
XX
CC The present invention relates to the identification of depression
CC associated protein isoforms (DPIs), particularly the tryptic digest
CC peptides of these proteins. Some of the DPIs (AAU28404-AAU28625)
CC described are decreased in the cerebrospinal fluid (CSF) of BAD
CC (bipolar affective disorder) subjects, whilst other DPIs
CC (AAU28626-AAU28887) are increased in BAD subjects. Also described
CC are peptide sequences identified from DPI-45 and DPI-213 and the
CC nucleic acid sequence they are encoded by. The sequences of the
CC invention are useful for clinical screening, diagnosis, prognosis,
CC therapy and prophylaxis of neuropsychiatric disorders e.g. BAD (also
CC known as bipolar mood disorder, BP), maniac-depressive illnesses,
CC attention deficit disorders, schizoaffective disorders, and unipolar
CC affective disorders. The present sequence represents one of the DPI
CC tryptic digest peptides of the present invention.
XX
SQ Sequence 8 AA;

Query Match 64.4%; Score 29; DB 22; Length 8;
Best Local Similarity 75.0%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NNATFYFK 8
||| |||
Db 1 nnahgyfk 8

RESULT 3
AAU24794
ID AAU24794 standard; Peptide; 8 AA.
XX
AC AAU24794;
XX
DT 18-DEC-2001 (first entry)
XX
DE Schizophrenia-Associated Protein Isoform (SPI) peptide #23.
XX
KW Schizophrenia-associated protein isoform; SPI; SPI-206; SPI-238; SPI-240;
KW neuroleptic; gene therapy; cerebrospinal fluid; serum; plasma.
XX
OS Homo sapiens.
XX
PN WO200162785-A2.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-GB00792.
XX
PR 24-FEB-2000; 2000GB-0004415.
PR 28-NOV-2000; 2000US-0750395.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Herath HMAC, Parekh RB, Rohlff C, Terrett JA, Tyson KL;
XX
DR WPI; 2001-570624/64.
XX
PT New schizophrenia associated protein isoforms and encoding nucleic acid
PT molecules, useful for treatment, diagnosis and prognosis of
PT schizophrenia and screening for potential drugs for treatment and new
PT drug targets -
XX
PS Disclosure; Page 29; 148pp; English.
XX
CC The sequence represents a schizophrenia-associated protein isoform (SPI).
CC These protein isoforms, e.g. SPI-206, SPI-238 and SPI-240 are detectable
CC in cerebrospinal fluid, serum or plasma and are useful markers of
CC schizophrenia. The sequences can be used for treatment and diagnosis of
CC schizophrenia, screening, prognosis, monitoring the results of therapy,
CC identifying patients most likely to respond to a particular therapy and
CC identification of new targets for drug treatment. SPI DNA is useful as a
CC nucleic acid probe to detect the presence of nucleic acids or SPIs.
XX
SQ Sequence 8 AA;

Query Match 64.4%; Score 29; DB 22; Length 8;
Best Local Similarity 75.0%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NNATFYFK 8
||| |||
Db 1 nnahgyfk 8

RESULT 4
AAU26242
ID AAU26242 standard; Peptide; 8 AA.
XX
AC AAU26242;
XX
DT 18-DEC-2001 (first entry)
XX
DE Depression-Associated Protein isoform DPI-205 #3.
XX
KW Human; Bipolar Affective Disorder; BAD; Depression-Associated feature;
KW DF; Depression-Associated protein isoform; DPI; Cerebro-spinal fluid;
KW CSF; antidepressant; antimanic; nootropic; tranquiliser; neuroleptic;
KW attention deficient disorder; schizoaffective disorder;
KW unipolar affective disorder.
XX
OS Homo sapiens.
XX
PN WO200163294-A2.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-GB00791.
XX
PR 24-FEB-2000; 2000GB-0004412.
PR 08-DEC-2000; 2000GB-0030050.
PR 12-DEC-2000; 2000US-0254830.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Herath HMAC, Parekh RB, Rohlff C;
XX
DR WPI; 2001-582081/65.
XX
PT Preparation for diagnosing or treating bipolar affected disorder (BAD)
PT or unipolar depression, or for screening for modulators, comprises a
PT BAD-associated protein isoform -
XX
PS Claim 8; Page 34; 163pp; English.
XX
CC The invention relates to a preparation comprising an isolated Bipolar
CC Affected Disorder (BAD)-Associated Protein Isoform (DPIs). The DPI's are
CC used to screen, diagnose or prognose of BAD or unipolar depression,
CC determine the stage or severity of BAD or unipolar depression, identify a
CC subject at risk of developing BAD or unipolar depression, or monitor the
CC effect of therapy in a subject. They are also used to screen for or
CC identify agents that interact with a DPI. These agents, antibodies
CC against the DPIs, and nucleic acids encoding the DPIs are used to treat
CC or prevent BAD or unipolar depression. Diseases that can be treated are
CC attention deficient disorder, a schizoaffective disorder, a bipolar or a
CC unipolar affective disorder. The DPIs are used in proteomics. The
CC proteomic approach of using DPIs for screening, diagnosis or prognosis of
CC BAD or unipolar depression overcomes the problems of using gene
CC expression analysis, such as not being able to obtain central nervous
CC system (CNS) tissue from a living patient under normal circumstances.
CC The present sequence is a DIP decreased in the CSF (cerebro-spinal
CC fluid) of subjects having BAD.
XX
SQ Sequence 8 AA;

Query Match 64.4%; Score 29; DB 22; Length 8;
Best Local Similarity 75.0%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NNATFYFK 8
||| |||
Db 1 nnahgyfk 8

RESULT 5
AAU15138
ID AAU15138 standard; Peptide; 8 AA.
XX
AC AAU15138;
XX
DT 24-OCT-2001 (first entry)
XX
DE Schizophrenia-associated isoform peptide #23.
XX
KW Schizophrenia; neuroleptic; diagnostic; neuropsychiatric disorder;
KW neurological disorder; neuropathy.
XX
OS Homo sapiens.
XX
PN WO200163293-A2.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-GB00783.
XX
PR 24-FEB-2000; 2000GB-0004415.
PR 28-NOV-2000; 2000US-0750395.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Herath HMAC, Parekh RB, Rohlff C;
XX
DR WPI; 2001-502868/55.
XX
PT Diagnosing and monitoring Schizophrenia by detecting the presence of
PT Schizophrenia Associated Features and Schizophrenia Associated Protein
PT Isoforms in samples of cerebrospinal fluid -
XX
PS Claim 6; Page 29; 160pp; English.
XX
CC The invention relates to methods and compositions for screening,
CC diagnosis and prognosis of Schizophrenia. The method involves detecting
CC the presence of Schizophrenia (SCH) Associated Features (SFs) and SCH
CC Associated Protein Isoforms (SPIs) in samples, e.g. by electrophoresis,
CC immunoassay or hybridisation assay, for diagnosing and monitoring SCH,
CC studying the effectiveness of treatments and for identifying potential
CC therapeutic agents. The method is used for (1) screening or diagnosis of
CC SCH and the relative abundance of at least 1 chosen feature correlates
CC with the presence or absence of SCH; and (2) monitoring the effect of
CC therapy administered to a subject with SCH and the relative abundance of
CC at least 1 chosen feature which correlates with the severity of SCH.
CC The expression and activity of the SFs, SPIs and related molecules
CC (e.g. secondary messengers) are studied to diagnose SCH, monitor the
CC progress of the disorder and the effectiveness of treatment and as
CC targets to identify and produce potential therapeutic agents for the
CC treatment of SCH. The paucity of detectable neuralgic defects
CC distinguishes neuropsychiatric disorders such as SCH from neurological
CC disorders, where manifestations of anatomical and biochemical changes
CC have been identified in many cases. Consequently the identification and
CC characterisation of cellular and/or molecular causative defects and
CC neuropathies are necessary for improved treatment of neuropsychiatric
CC disorders. AAU15114-AAU15762 represent the amino acid sequences of
CC schizophrenia-associated isoforms used in the method of the invention.
XX
SQ Sequence 8 AA;

Query Match 64.4%; Score 29; DB 22; Length 8;
Best Local Similarity 75.0%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 NNATFYFK 8
          |||  |||
Db      1 nnahgyfk 8


RESULT   6
AAR73374
ID   AAR73374 standard; Peptide; 8 AA.
XX
AC   AAR73374;
XX
DT   08-DEC-1995  (first entry)
XX
DE   Human TSH receptor (residues 347-354).
XX
KW   thyroid stimulating hormone receptor; TSH; human; Homo sapiens;
KW   antibody; affinity; detection.
XX
OS   Synthetic.
XX
PN   JP07089991-A.
XX
PD   04-APR-1995.
XX
PF   28-SEP-1993;   93JP-0240853.
XX
PR   28-SEP-1993;   93JP-0240853.
XX
PA   (MITP ) MITSUBISHI PETROCHEMICAL CO LTD.
XX
DR   WPI; 1995-167251/22.
XX
PT   Novel polypeptide(s) having affinity for the human TSH receptor
PT   antibody - used in detection of the TSH antibody.
XX
PS   Example 1; Page 28; 54pp; Japanese.
XX
CC   Peptides with affinity to human TSH (thyroid stimulating hormone)
CC   receptor antibody are used for detection of the antibody. (See also
CC   AAR73201-592).
XX
SQ   Sequence   8 AA;


  Query Match            53.3%;  Score 24;  DB 16;  Length 8;
  Best Local Similarity  66.7%;  Pred. No. 6.4e+05;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 NNATFY 6
          ||| :|
Db      2 nnahyy 7


RESULT   7
AAU08641
ID   AAU08641 standard; Peptide; 6 AA.
XX
AC   AAU08641;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Interleukin-1 inhibitor based peptide #5.
XX
KW   Interleukin-1 inhibitor; IL-1i; antiarthritic;
KW   antirheumatic; antiulcer; osteopathic; antidiabetic; antipsoriatic;
KW   immunosuppressive; antigout; antiinflammatory; rheumatoid arthritis;
KW   Crohn's disease; ulcerative colitis; osteoporosis; juvenile diabetes;
KW   psoriasis; lupus erythematosus; fibrosis; glomerulonephritis; gout;
KW   acute febrile illness; sarcoidosis; lymphoma; vascular occlusion;
KW   interstitial lung disease.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1..6
FT                   /note= "Encoded by AAYCARAARACNTTYTA"
XX
PN   EP1133995-A2.
XX
PD   19-SEP-2001.
XX
PF   26-MAY-1989; 2001EP-0107950.
XX
PR   27-MAY-1988;   88US-0199915.
PR   31-AUG-1988;   88US-0238713.
PR   23-SEP-1988;   88US-0248521.
PR   03-NOV-1988;   88US-0266531.
PR   26-MAY-1989;   89EP-0109540.
PR   26-MAY-1989;   92EP-0114773.
XX
PA   (SYND ) SYNERGEN INC.
XX
PI   Hannum CH,  Eisenberg SP,  Thompson RC,  Arend WP,  Joslin FG;
PI   Sommer A;
XX
DR   WPI; 2001-566880/64.
DR   N-PSDB; AAS13435.
XX
PT   Recombinant interleukin-1 inhibitors and polynucleotides encoding the
PT   protein useful for treating an IL-1 mediated pathophysiological
PT   condition which includes arthritis, rheumatoid arthritis, Crohn's
PT   disease -
XX
PS   Example 6; Page 22; 58pp; English.
XX
CC   The invention relates to a recombinant polypeptide having interleukin-1
CC   inhibitor (IL-1i) activity against IL-1alpha, IL-1beta or both.
CC   The IL-1i is useful for preparing pharmaceutical composition for
CC   treating an IL-1 mediated pathophysiological condition which includes
CC   arthritis, rheumatoid arthritis, Crohn's disease, ulcerative colitis,
CC   osteoporosis, juvenile diabetes, psoriasis, lupus erythematosus,
CC   fibrosis, glomerulonephritis, gout, acute febrile illness, sarcoidosis,
CC   lymphomas, a patient after damage to the brain from vascular occlusion,
CC   or acute or chronic interstitial lung disease. The present sequence
CC   is a peptide based on an IL-1i of the invention, used to design
CC   degenerate probes for isolating nucleic acids encoding IL-1i molecules.
XX
SQ   Sequence   6 AA;


  Query Match            51.1%;  Score 23;  DB 22;  Length 6;
  Best Local Similarity  66.7%;  Pred. No. 6.4e+05;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 NNATFY 6
          |  |||
Db      1 nqktfy 6


RESULT   8
AAY40730
ID   AAY40730 standard; peptide; 7 AA.
XX
AC   AAY40730;
XX
DT   01-DEC-1999  (first entry)
XX
DE   S4 derivative #4, beta strand of scaffold protein structure.
XX
KW   Scaffold protein; beta strand; beta sheet; stabilize antigen; vaccine;
KW   tumour; chemotherapeutic agent.
XX
OS   Synthetic.
XX
PN   EP947582-A1.
```

XX
PD 06-OCT-1999.
XX
PF 31-MAR-1998; 98EP-0870065.
XX
PR 31-MAR-1998; 98EP-0870065.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Desmet J, Hufton S, Hoogenboom H, Sablon E;
XX
DR WPI; 1999-542958/46.
XX
PT New scaffold protein, useful for stabilizing antigens used as vaccines
PT -
XX
PS Disclosure; Page 6; 105pp; English.
XX
CC Sequences AAY40727-Y40748 are functionally equivalent derivatives of the
CC S4 peptide (AAY40607) which forms part of a scaffold protein. S4 is a
CC beta strand peptide which forms part of a beta sheet. Peptides
CC (AAY40601-Y40609) together form a single-chain scaffold protein which
CC contains at least 1 disulfide bond, contains less than 10% alpha helix
CC and contains at least 6 beta-strands. The scaffold protein is constructed
CC of beta strands S1-S6, and may also include beta strands A1-A3, or any
CC functionally equivalent derivative of these sequences. The beta strands
CC form two beta sheets S1/S4/S3 and S6/S5/S2 with each strand connected to
CC the next by hydrogen bonds, which generate a beta sandwich architecture.
CC If the additional beta strands A1-A3 are included in the structure the
CC scaffold is constructed of two beta sheets, with the structures
CC A1/S1/S4/S3 and S6/S5/S2/A2/A3. The beta strands are connected to each
CC other via amino acid loops, where at least one of the loops binds to a
CC receptor or antigen. The scaffold protein is used to stabilize antigens
CC or whole proteins such as receptors, or their fragments. It may be used
CC to bind two separate molecules. For example, one surface of the scaffold
CC may be bound to a protein which binds to a tumour antigen. This will
CC target the complex to tumour cells. Another surface may be bound to a
CC cytotoxic molecule or an autoimmune antibody which may then kill the
CC tumour cells. Therefore the scaffold protein may be used to target
CC chemotherapeutic agents to specific cells. It may also be used to
CC stabilize individual peptides in a peptide library and may be used in
CC diagnostic techniques, and to stabilize antigens used as vaccines.
XX
SQ Sequence 7 AA;

Query Match 51.1%; Score 23; DB 20; Length 7;
Best Local Similarity 80.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 TFYFK 8
||| |
Db 3 tfylk 7

RESULT 9
AAB30068
ID AAB30068 standard; Peptide; 7 AA.
XX
AC AAB30068;
XX
DT 09-FEB-2001 (first entry)
XX
DE Scaffold protein SCA S4 peptide SEQ ID NO: 129.
XX
KW Human; CTLA-4; scaffold protein; antigen-binding; receptor-binding;
KW SCA domain; cancer; thrombosis; osteoporosis, rheumatoid arthritis;
KW diabetic retinopathy; atherosclerosis.
XX
OS Synthetic.
XX
PN WO200060070-A1.
XX
PD 12-OCT-2000.
XX
PF 01-APR-1999; 99WO-EP02283.
XX
PR 01-APR-1999; 99WO-EP02283.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Desmet J, Hufton S, Hoogenboom H, Sablon E;
XX
DR WPI; 2000-665002/64.
XX
PT Scaffold composed of single-chain polypeptide having beta sandwich
PT architecture carrying new and randomized peptide sequences useful as
PT supporting framework and carrying antigen- or receptor binding
PT fragments -
XX
PS Disclosure; Page 15; 68pp; English.
XX
CC The present invention is concerned with producing scaffold proteins
CC based upon the human CTLA-4 SCA domain. These scaffold proteins can be
CC used as a scaffold to bind antigen- or receptor-binding fragments. These
CC can be used in the treatment of diseases such as cancer,
CC atherosclerosis, thrombosis, osteoporosis, rheumatoid arthritis and
CC diabetic retinopathy. Sequences AAB29930-B29939 were used in the
CC production of the proteins of the invention.
XX
SQ Sequence 7 AA;

Query Match 51.1%; Score 23; DB 21; Length 7;
Best Local Similarity 80.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 TFYFK 8
||| |
Db 3 tfylk 7

RESULT 10
AAY23893
ID AAY23893 standard; Peptide; 6 AA.
XX
AC AAY23893;
XX
DT 21-SEP-1999 (first entry)
XX
DE Human vascular endothelial growth factor (VEGF)-D derived peptide.
XX
KW Vascular endothelial growth factor; VEGF; VEGF-D; malignant melanoma;
KW tumour; psoriasis; angiogenesis; lymphangiogenesis; skin graft;
KW wound healing; lymphedema; scleroderma; anhydrotic ectodermal dysplasia.
XX
OS Homo sapiens.
XX
PN WO9933485-A1.
XX
PD 08-JUL-1999.
XX
PF 23-DEC-1998; 98WO-US27373.
XX
PR 29-MAY-1998; 98US-0087392.
PR 24-DEC-1997; 97AU-0001131.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Achen MG, Alitalo K, Stacker SA;
XX
DR WPI; 1999-405368/34.
XX
PT A human cell line stably expressing vascular endothelial growth
PT factor D, useful for treating melanomas or tumours expressing VEGF-D
XX

PS Example 7; Page 75; 79pp; English.
XX
CC The present sequence is derived from vascular endothelial growth
CC factor (VEGF)-D. The specification describes a human cell line
CC which stably expresses VEGF-D, or fragments/analogues having VEGF-D
CC biological activity. VEGF-D antagonists, e.g. antisense nucleic acids
CC or triplex DNA, VEGF-D variants or antibodies (especially chimeric
CC antibodies), are useful for the treatment or alleviation of malignant
CC melanomas, tumours or psoriasis. Angiogenesis and lymphangiogenesis
CC stimulating amounts of VEGF-D can be administered to enhance the
CC acceptance and/or healing of skin grafts or to stimulate the healing
CC of a surgical or traumatic wound to the skin. Lymphangiogenesis
CC stimulating amounts of VEGF-D can be used to treat lymphedema.
CC Endothelial proliferation stimulating amounts of VEGF-D are used to
CC treat scleroderma. Vascularisation stimulating amounts of VEGF-D can
CC be used to treat anhydrotic ectodermal dysplasia. VEGF-D antibodies are
CC useful for detecting tumours expressing VEGF-D. Fully-processed VEGF-D
CC can be used to stimulate at least one VEGF-D bioactivity chosen from
CC endothelial cell proliferation, migration, survival and differentiation
CC and lymphangiogenesis without inducing vascular permeability.
XX
SQ Sequence 6 AA;

Query Match 48.9%; Score 22; DB 20; Length 6;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 ATFY 6
||||
Db 3 atfy 6

RESULT 11
AAW97492
ID AAW97492 standard; peptide; 6 AA.
XX
AC AAW97492;
XX
DT 19-MAY-1999 (first entry)
XX
DE Antigenic site of HN protein loop beta-3L01.
XX
KW Antigenic site; haemagglutinin-neuraminidase; HN; paramyxoviridae;
KW virus epitope; attachment protein; vaccine; immunodominant epitope.
XX
OS Bovine parainfluenza virus.
XX
PN WO9902695-A2.
XX
PD 21-JAN-1999.
XX
PF 08-JUL-1998; 98WO-NL00390.
XX
PR 08-JUL-1997; 97EP-0202100.
XX
PA (DIER-) STICHTING INST DIERHOUDERIJ EN DIERGEZON.
XX
PI Langedijk JPM, Van Oirschot JT;
XX
DR WPI; 1999-120896/10.
XX
PT Isolated proteinaceous substance - comprising at least one virus
PT epitope derived from an attachment protein of a paramyxovirus
XX
PS Disclosure; Page 45; 63pp; English.
XX
CC AAW97452-571 represent antigenic sites derived from the
CC haemagglutinin-neuraminidase (HN) protein of the paramyxoviridae.
CC The specification describes 3-D models identifying a proteinaceous
CC substance comprising at least one virus epitope derived from the
CC attachment protein, which corresponds to an antigenic site present on
CC one of the loops of HN. The antigenic sites can be used to produce
CC vaccines, to detect the viruses, and to select the immunodominant
CC epitope.
XX
SQ Sequence 6 AA;

Query Match 48.9%; Score 22; DB 20; Length 6;
Best Local Similarity 80.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NNATF 5
|| ||
Db 1 nnitf 5

RESULT 12
AAY23895
ID AAY23895 standard; Peptide; 7 AA.
XX
AC AAY23895;
XX
DT 21-SEP-1999 (first entry)
XX
DE Human vascular endothelial growth factor (VEGF)-D derived peptide.
XX
KW Vascular endothelial growth factor; VEGF; VEGF-D; malignant melanoma;
KW tumour; psoriasis; angiogenesis; lymphangiogenesis; skin graft;
KW wound healing; lymphedema; scleroderma; anhydrotic ectodermal dysplasia.
XX
OS Homo sapiens.
XX
PN WO9933485-A1.
XX
PD 08-JUL-1999.
XX
PF 23-DEC-1998; 98WO-US27373.
XX
PR 29-MAY-1998; 98US-0087392.
PR 24-DEC-1997; 97AU-0001131.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Achen MG, Alitalo K, Stacker SA;
XX
DR WPI; 1999-405368/34.
XX
PT A human cell line stably expressing vascular endothelial growth
PT factor D, useful for treating melanomas or tumours expressing VEGF-D
XX
PS Example 7; Page 75; 79pp; English.
XX
CC The present sequence is derived from vascular endothelial growth
CC factor (VEGF)-D. The specification describes a human cell line
CC which stably expresses VEGF-D, or fragments/analogues having VEGF-D
CC biological activity. VEGF-D antagonists, e.g. antisense nucleic acids
CC or triplex DNA, VEGF-D variants or antibodies (especially chimeric
CC antibodies), are useful for the treatment or alleviation of malignant
CC melanomas, tumours or psoriasis. Angiogenesis and lymphangiogenesis
CC stimulating amounts of VEGF-D can be administered to enhance the
CC acceptance and/or healing of skin grafts or to stimulate the healing
CC of a surgical or traumatic wound to the skin. Lymphangiogenesis
CC stimulating amounts of VEGF-D can be used to treat lymphedema.
CC Endothelial proliferation stimulating amounts of VEGF-D are used to
CC treat scleroderma. Vascularisation stimulating amounts of VEGF-D can
CC be used to treat anhydrotic ectodermal dysplasia. VEGF-D antibodies are
CC useful for detecting tumours expressing VEGF-D. Fully-processed VEGF-D
CC can be used to stimulate at least one VEGF-D bioactivity chosen from
CC endothelial cell proliferation, migration, survival and differentiation
CC and lymphangiogenesis without inducing vascular permeability.
XX
SQ Sequence 7 AA;

```
  Query Match             48.9%;  Score 22;  DB 20;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 6.4e+05;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 ATFY 6
           ||||
Db       4 atfy 7


RESULT  13
AAW20072
ID   AAW20072 standard; peptide; 4 AA.
XX
AC   AAW20072;
XX
DT   12-SEP-1997  (first entry)
XX
DE   HOX5, a hexose oxidase derived antimicrobial/antioxidant peptide.
XX
KW   Hexose oxidase; Chondrus crispus; marine algae; recombinant;
KW   antimicrobial; antioxidant; food preparation; dairy product; starch;
KW   beverage; animal feed; silage; sugar reduction; cosmetics; dental;
KW   toothpaste; dough; lactone production.
XX
OS   Chondrus crispus.
XX
PN   WO9640935-A1.
XX
PD   19-DEC-1996.
XX
PF   04-JUN-1996;   96WO-DK00238.
XX
PR   07-JUN-1995;   95US-0476910.
XX
PA   (BIOT-) BIOTEKNOLOGISK INST.
XX
PI   Hansen OC,  Stougaard P;
XX
DR   WPI; 1997-052332/05.
XX
PT   Recombinant production of polypeptide having hexose oxidase activity
PT   - used in food preparations as antibacterial and antioxidant agent
XX
PS   Claim 6; Page 100; 124pp; English.
XX
CC   AAW20068-75 are small peptides derived from hexose oxidase (HO) of the
CC   marine algae species Chondrus crispus. Hexose oxidase and peptides
CC   having HO activity are useful in the production of food products, e.g.
CC   dairy products, starch-containing food products (dough) and non-dairy
CC   beverages. HO and active peptide fragments have antimicrobial and
CC   antioxidant properties and act by removing all the oxygen in a food
CC   packaging. HO and peptides of HO can also be used in an animal feed,
CC   especially silage. Further uses are to reduce, or analyse, the sugar
CC   content in a food, in the production of cosmetics, tooth care products
CC   or a pharmaceutical product and in lactone production. HO can be
CC   recombinantly produced in industrially appropriate quantities, and at
CC   a quality and purity level which renders polypeptide suitable for
CC   industrial purposes.
XX
SQ   Sequence   4 AA;


  Query Match             46.7%;  Score 21;  DB 18;  Length 4;
  Best Local Similarity   75.0%;  Pred. No. 6.4e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 FYFK 8
           :|||
Db       1 yyfk 4


RESULT  14
AAW00431
ID   AAW00431 standard; peptide; 5 AA.
XX
AC   AAW00431;
XX
DT   29-AUG-1996  (first entry)
XX
DE   Interleukin-6 antagonist peptide.
XX
KW   IL-6; antagonist; autoimmune disease.
XX
OS   Synthetic.
XX
PN   JP07324097-A.
XX
PD   12-DEC-1995.
XX
PF   30-MAY-1994;   94JP-0117259.
XX
PR   30-MAY-1994;   94JP-0117259.
XX
PA   (DAIL ) DAICEL CHEM IND LTD.
PA   (FUJI ) FUJISAWA PHARM CO LTD.
XX
DR   WPI; 1996-065476/07.
XX
PT   Interleukin 6 antagonist - useful for treating auto:immune diseases
XX
PS   Claims 4, 6; Pages 2, 3; 19pp; Japanese.
XX
CC   New IL-6 antagonists are provided which are of formula X-W-Y, in
CC   which X is H or an amino-protecting group, Y is OH or a carboxy-
CC   protecting group, and W is a peptide containing all or part of the
CC   sequence as given in AAW00401, AAW00402, AAW00403 or AAW00404, where any
CC   free mercapto groups in the sequence are optionally protected. The
CC   present sequence is a specifically preferred partial sequence of AAW00403
CC   and is itself claimed as a new chemical entity.
CC   The IL-6 antagonists are useful for treating autoimmune diseases.
XX
SQ   Sequence   5 AA;


  Query Match             46.7%;  Score 21;  DB 17;  Length 5;
  Best Local Similarity   60.0%;  Pred. No. 6.4e+05;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 NATFY 6
           |::||
Db       1 nssfy 5


RESULT  15
AAB04166
ID   AAB04166 standard; Peptide; 6 AA.
XX
AC   AAB04166;
XX
DT   11-APR-2001  (first entry)
XX
DE   Internal peptide fragment of fructosyl transferase.
XX
KW   Fructan exohydrolase; FEH; transgenic plant; recombination;
KW   transgene; gene expression; detergent; detergen additive;
KW   oral care composition.
XX
OS   Cichorium intybus.
XX
PN   WO200068402-A1.
XX
PD   16-NOV-2000.
XX
PF   08-MAY-2000; 2000WO-EP04226.
XX
PR   06-MAY-1999;   99BE-0000329.
```

```
XX
PA   (LEUV-) LEUVEN RES & DEV.
XX
PI   Van Den Ende W,  Van Laere A,  De Roover J,  Michiels A;
XX
DR   WPI; 2001-007401/01.
XX
PT   Novel DNA molecules encoding enzymes having fructan exohydrolase
PT   activity for use in transgenic plant production, dental care
PT   compositions, and in detergents
XX
PS   Example 1; Page 16; 45pp; English.
XX
CC   Transgenic plants such as Cichorium intybus, Cynara scolymus,
CC   Helianthus tuberosus, Scorzonera hispanica, Oryza sativa, Zea mays,
CC   Triticum aestivum, Triticum durum, Hordeum vulgare, Secale cereale,
CC   Avena sativa, Sorghum vulgare, Phleum pratense, Lolium temulentum,
CC   Dactylis glomerata, Pennisetum americanum, Allium cepa, Agave
CC   americanum, Agava azul teqilana, Sorghum bicolor and Panicum milaceum,
CC   transformed with a vector encoding a fructan exohydrolase (FEH)
CC   enzyme are useful for the recombinant production of FEH or other
CC   polypeptides having FEH activity. The FEH polypeptides produced
CC   are useful in detergents or as a detergent additive and in oral
CC   care compositions. This internal conserved peptide sequence of
CC   fructosyl transferase and invertase was used to create a
CC   degenerate antisense primer for amplifying the FEH coding sequence.
XX
SQ   Sequence   6 AA;


  Query Match             46.7%;  Score 21;  DB 22;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 6.4e+05;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 NNAT 4
          ||||
Db      2 nnat 5


RESULT  16
AAW11529
ID   AAW11529 standard; peptide; 8 AA.
XX
AC   AAW11529;
XX
DT   22-APR-1997  (first entry)
XX
DE   Interleukin-6 antagonist 2.
XX
KW   Interleukin-6; IL-6; antagonist; inhibitor; autoimmune disease;
KW   skin; intestine; systemic lupus erythematosus; chronic rheumatism.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   8
FT                   /note= "amidated"
XX
PN   JP08311098-A.
XX
PD   26-NOV-1996.
XX
PF   22-MAY-1995;   95JP-0146742.
XX
PR   22-MAY-1995;   95JP-0146742.
XX
PA   (DAIL ) DAICEL CHEM IND LTD.
PA   (FUJI ) FUJISAWA PHARM CO LTD.
XX
DR   WPI; 1997-061811/06.
XX
PT   Interleukin-6 antagonistic peptide(s) comprising arginine - useful
PT   for treating autoimmune, renal, skin and intestinal diseases
XX
PS   Example 2; Page 11; 20pp; Japanese.
XX
CC   The present peptide is a specific example of new interleukin-6
CC   antagonists of the general formula X-A-B-D-Y, where X is 1-10 amino
CC   acids or an amino group protecting group; Y is 1-5 amino acids, a
CC   carboxyl group protecting group or an amide; A is preferably Arg
CC   having an opt. protected guanidino group but can be any amino acid;
CC   D is Arg having an opt. protected guanidino group and B is preferably
CC   a Leu residue but can be any amino acid, including non-natural
CC   amino acids, opt. having a protected side-chain. The peptides are
CC   useful for treating autoimmune diseases (e.g. systemic lupus
CC   erythematosus or chronic rheumatism), renal, skin and intestinal
CC   diseases.
XX
SQ   Sequence   8 AA;


  Query Match             46.7%;  Score 21;  DB 18;  Length 8;
  Best Local Similarity   60.0%;  Pred. No. 6.4e+05;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 NATFY 6
            |::||
Db      1 nssfy 5


RESULT  17
AAW86298
ID   AAW86298 standard; peptide; 5 AA.
XX
AC   AAW86298;
XX
DT   19-FEB-1999  (first entry)
XX
DE   Rodent IL-1 delta polypeptide fragment (residues 150-154).
XX
KW   Interleukin; IL-1 delta; polyclonal antibody; IL-1 epsilon; cytokine;
KW   inflammatory response; immune system; diagnosis; agonist; antagonist;
KW   chemokine.
XX
OS   Mus sp.
XX
PN   WO9847921-A1.
XX
PD   29-OCT-1998.
XX
PF   17-APR-1998;   98WO-US06879.
XX
PR   06-AUG-1997;   97US-0055111.
PR   21-APR-1997;   97US-0837627.
XX
PA   (SCHE ) SCHERING CORP.
XX
PI   Bazan JF,  Hedrick JA,  Kastelein RA,  Sana TR;
XX
DR   WPI; 1998-609976/51.
XX
PT   Mammalian interleukin 1-delta and 1-epsilon - useful for, e.g.
PT   regulating the immune system and inflammatory responses
XX
PS   Claim 1; Page 99; 113pp; English.
XX
CC   The invention relates to a recombinant polypeptide that specifically
CC   binds polyclonal antibodies (Abs) generated against a 12 consecutive
CC   amino acid segment of interleukin (IL)-1 delta or IL-1 epsilon. Agonists
CC   or antagonists of these IL polypeptides are used to regulate a cell
CC   involved in an inflammatory response. The IL-1 delta or IL-1 epsilon
CC   polypeptides and peptides are used to produce Abs and antigen-Ab
CC   complexes. The polypeptides, Abs and the corresponding nucleic acids
CC   regulate development and/or the immune system, and can be used to
CC   diagnose and treat conditions associated with abnormal expression of IL.
CC   Agonists or antagonists of IL-1 delta or IL-1 epsilon polypeptides are
```

```
CC   used with agonists or antagonists of IL-1 alpha, IL-1RA, IL-1 beta, IL-1
CC   gamma, IL-2 and/or IL-12. The IL-1 delta or IL-1 epsilon polypeptides
CC   may be used as a soluble polypeptide or as a fusion protein with another
CC   cytokine or chemokine. Sequences AAW86287 to AAW86300 represent peptide
CC   fragments of a rodent interleukin (IL)-1 delta polypeptide.
XX
SQ   Sequence   5 AA;


  Query Match             44.4%;  Score 20;  DB 19;  Length 5;
  Best Local Similarity   75.0%;  Pred. No. 6.4e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 FYFK 8
           |||:
Db       1 fyfq 4


RESULT  18
AAY25842
ID   AAY25842 standard; Protein; 6 AA.
XX
AC   AAY25842;
XX
DT   04-OCT-1999  (first entry)
XX
DE   Human secreted protein fragment encoded from gene 49.
XX
KW   Secreted protein; human; treatment; diagnosis; therapy; cancer; tumour;
KW   neurodegenerative disorder; developmental abnormality; blood disorder;
KW   fetal deficiency; blood disorder; leukemia; immune system; inflammation;
KW   autoimmune disease; hepatic disease; renal disease; allergy; restenosis;
KW   ischaemic shock; Alzheimer's disease; cognitive disorder; schizophrenia;
KW   cardiovascular disorder; wound healing; stroke; arthritis; obesity;
KW   asthma; sepsis; acne; psoriasis; transplant rejection; infection; AIDS;
KW   metabolic disorder.
XX
OS   Homo sapiens.
XX
PN   WO9938881-A1.
XX
PD   05-AUG-1999.
XX
PF   27-JAN-1999;   99WO-US01621.
XX
PR   30-JAN-1998;   98US-0073170.
PR   30-JAN-1998;   98US-0073159.
PR   30-JAN-1998;   98US-0073160.
PR   30-JAN-1998;   98US-0073161.
PR   30-JAN-1998;   98US-0073162.
PR   30-JAN-1998;   98US-0073164.
PR   30-JAN-1998;   98US-0073165.
PR   30-JAN-1998;   98US-0073167.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Carter KC,  Endress GA,  Feng P,  Ferrie AM,  Florence C;
PI   Florence KA,  Janat F,  Ni J,  Rosen CA,  Ruben SM;
PI   Soppet DR,  Young P,  Yu G;
XX
DR   WPI; 1999-469315/39.
DR   N-PSDB; AAZ00458.
XX
PT   New isolated human genes and the secreted polypeptides they encode
PT   useful in, e.g. treatment of Alzheimer's
XX
PS   Disclosure; Page 365; 393pp; English.
XX
CC   This invention describes novel human genes (see AAZ00410-Z00477) and the
CC   secreted proteins (see AAY25711-Y25778) and fragments (see
CC   AAY25779-Y25907) they encode. The polynucleotides and their corresponding
CC   secreted polypeptides are useful for preventing, treating or ameliorating
CC   medical conditions e.g. by protein or gene therapy. Also pathological
CC   conditions can be diagnosed by determining the amount of the new
CC   polypeptides in a sample or by determining the presence of mutations in
CC   the new polynucleotides. Specific uses are described for each of the 67
CC   polynucleotides of the invention, based on which tissues they are most
CC   highly expressed in, and include developing products for the diagnosis
CC   or treatment of cancer, tumours, neurodegenerative disorders,
CC   developmental abnormalities and fetal deficiencies, blood disorders,
CC   leukemias, diseases of the immune system, autoimmune diseases, hepatic
CC   and renal disease, inflammation, allergies, ischaemic shock, Alzheimer's
CC   and cognitive disorders, schizophrenia, restenosis, cardiovascular
CC   disorders, wound healing, stroke, arthritis, obesity, asthma, sepsis,
CC   acne, psoriasis, transplant rejection, metabolic disorders, infections
CC   and AIDS. The polypeptides are also useful for identifying their binding
CC   partners.
XX
SQ   Sequence   6 AA;


  Query Match             44.4%;  Score 20;  DB 20;  Length 6;
  Best Local Similarity   66.7%;  Pred. No. 6.4e+05;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 ATFYFK 8
           |  |||
Db       1 arvyfk 6


RESULT  19
AAU08625
ID   AAU08625 standard; Peptide; 6 AA.
XX
AC   AAU08625;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Interleukin-1 inhibitor IL-1i-a (RalphaLysC-31).
XX
KW   Interleukin-1 inhibitor; IL-1i-a; RalphaLysC-31; antiarthritic;
KW   antirheumatic; antiulcer; osteopathic; antidiabetic; antipsoriatic;
KW   immunosuppressive; antigout; antiinflammatory; rheumatoid arthritis;
KW   Crohn's disease; ulcerative colitis; osteoporosis; juvenile diabetes;
KW   psoriasis; lupus erythematosus; fibrosis; glomerulonephritis; gout;
KW   acute febrile illness; sarcoidosis; lymphoma; vascular occlusion;
KW   interstitial lung disease.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   EP1133995-A2.
XX
PD   19-SEP-2001.
XX
PF   26-MAY-1989; 2001EP-0107950.
XX
PR   27-MAY-1988;   88US-0199915.
PR   31-AUG-1988;   88US-0238713.
PR   23-SEP-1988;   88US-0248521.
PR   03-NOV-1988;   88US-0266531.
PR   26-MAY-1989;   89EP-0109540.
PR   26-MAY-1989;   92EP-0114773.
XX
PA   (SYND ) SYNERGEN INC.
XX
PI   Hannum CH,  Eisenberg SP,  Thompson RC,  Arend WP,  Joslin FG;
PI   Sommer A;
XX
DR   WPI; 2001-566880/64.
XX
PT   Recombinant interleukin-1 inhibitors and polynucleotides encoding the
PT   protein useful for treating an IL-1 mediated pathophysiological
PT   condition which includes arthritis, rheumatoid arthritis, Crohn's
PT   disease -
XX
```

PS Example 3; Page 19; 58pp; English.
XX
CC The invention relates to a recombinant polypeptide having interleukin-1
CC inhibitor (IL-1i) activity against IL-1alpha, IL-1beta or both.
CC The IL-1i is useful for preparing pharmaceutical composition for
CC treating an IL-1 mediated pathophysiological condition which includes
CC arthritis, rheumatoid arthritis, Crohn's disease, ulcerative colitis,
CC osteoporosis, juvenile diabetes, psoriasis, lupus erythematosus,
CC fibrosis, glomerulonephritis, gout, acute febrile illness, sarcoidosis,
CC lymphomas, a patient after damage to the brain from vascular occlusion,
CC or acute or chronic interstitial lung disease. The present sequence
CC is an IL-1i of the invention.
XX
SQ Sequence 6 AA;

Query Match 44.4%; Score 20; DB 22; Length 6;
Best Local Similarity 75.0%; Pred. No. 6.4e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FYFK 8
|||:
Db 1 fyfq 4

RESULT 20
AAW94661
ID AAW94661 standard; peptide; 7 AA.
XX
AC AAW94661;
XX
DT 30-APR-1999 (first entry)
XX
DE GL-7ACA acylase alternative N-terminal peptide #3.
XX
KW 7-beta-(4-carboxybutanamido)-cephalosporanic acid acylase;
KW GL-7ACA acylase; 7-amino-cephalosporin; cephalosporin.
XX
OS Synthetic.
XX
PN JP07313161-A.
XX
PD 05-DEC-1995.
XX
PF 24-MAY-1994; 94JP-0110066.
XX
PR 24-MAY-1994; 94JP-0110066.
XX
PA (FUJI) FUJISAWA PHARM CO LTD.
XX
DR WPI; 1996-053784/06.
XX
PT New carboxy:butanamido-cephalosporanic acid acylase - prepd by
PT recombinant DNA methods, used for prepn of 7-amino-cephalosporin
PT cpds
XX
PS Claim 1; Page 2; 18pp; Japanese.
XX
CC The present invention describes 7-beta-(4-carboxybutanamido)-
CC cephalosporanic acid acylase (GL-7ACA acylase) which has an amino acid
CC sequence in which part of the N-terminal sequence of natural GL-7ACA
CC acylase is replaced by at least one amino acid sequence selected from
CC AAW94659 to AAW94663. The GL-7ACA acylases of the present invention can
CC be prepared in a large amount and they are useful for the commercial
CC production of cephalosporin compounds.
XX
SQ Sequence 7 AA;

Query Match 44.4%; Score 20; DB 17; Length 7;
Best Local Similarity 75.0%; Pred. No. 6.4e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FYFK 8
|||:
Db 2 fyfe 5

Search completed: July 8, 2002, 11:38:27
Job time: 39 sec

US 094610610WP1

Creation date: 02-10-2004
Indexing Officer: THAILEMARIAM - TIZITA HAILEMARIAM
Team: OIPEBackFileIndexing
Dossier: 09461061

Legal Date: 07-10-2002

No.	Doccode	Number of pages
1	SRNT	40

Total number of pages: 40

Remarks:

Order of re-scan issued on .....